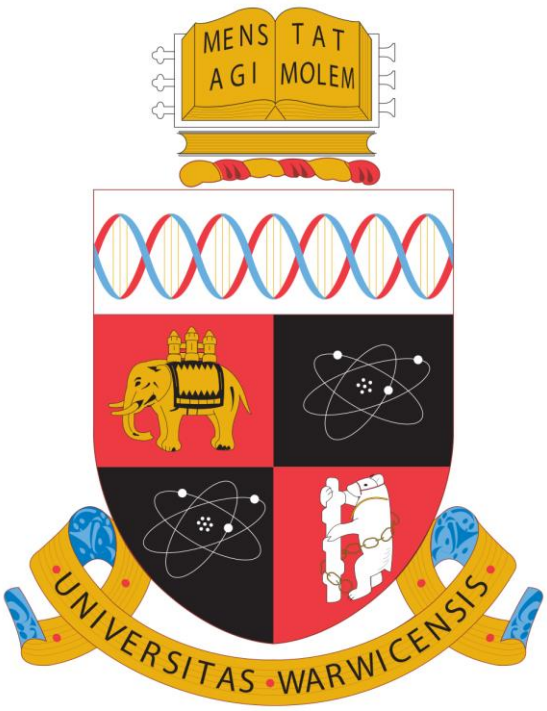


Metallodrug Interactions with Proteins in Biological Systems Studied by Fourier Transform-Ion Cyclotron Resonance Mass Spectrometry

Kung Ching Cookson Chiu

A thesis submitted for the degree of

Doctor of Philosophy

Department of Chemistry

University of Warwick

September 2019

Acknowledgements

Foremost, I would like to offer my sincerest gratitude to my two supervisors, Professor Peter B. O'Connor and Professor Peter J. Sadler, for giving me this valuable opportunity to undertake my PhD study. It is my pleasure to be their PhD student and work so closely with these two top scientists in the world.

It is my greatest honour to have Peter O'Connor as my mass spectrometry mentor. Pete's enthusiasm to the mass spectrometry field has always been remarkable. His in-depth knowledge of the research field and all the support he has selflessly given to me throughout the four years of my PhD study has been the steady guiding force during my research. Pete has always reminded us to analyse all the data critically and only draw conclusion from the analysis instead of the expectation. Pete also encourages us to attend as many conferences as possible, which helps to broaden my horizons in different research areas and learn from other world experts in the field. It is my honour to be his student for both my master and PhD studies, and I am looking forward to meeting him in conferences around the world and maybe one day collaborating with him again in the future!

I would like to extend my gratitude to Peter Sadler for his limitless support and trust in the four years of my PhD study. Peter has always been fantastic to talk to and his abundant knowledge in different areas in Chemistry is nothing short of amazing. Peter has always been extremely supportive of my findings, even if they were controversial. His trust in me and my data, and guidance throughout my PhD has been priceless and mean a lot to me. Without his support and advice, none of this work presented within would have been possible.

I would like to acknowledge all the group members, past and present, in the O'Connor group, the Sadler group, and the Barrow group. I have learnt so much from them, not only from the academic side, but also the interesting cultures, and different viewpoints. Working alongside these passionate scientists have made my PhD journey so much more enjoyable and interesting. I would like to specially acknowledge Dr. Mark P. Barrow, Dr. Christopher A. Wootton, Dr. Yuko P. Y. Lam, and Miss Meng Li

for the help, support, and fun during my PhD. I would also thank all the co-authors and collaborators who have been involved in the projects presented in this thesis.

I would like to take this opportunity to thank the Warwick University Snooker and Billiards Club, my parents, my family, and my partner Yvonne, for their support and advice especially in the difficult times during my PhD. Without your support, I would never be able to finish this milestone of my life.

Thank you everyone!

Author Information and *Curriculum Vitae*

Cookson Kung Ching Chiu

Email: [redacted]

Education

Doctor of Philosophy, Chemistry

University of Warwick October 2015 – September 2019

Thesis title: Metallodrug Interactions with Proteins in Biological Systems Studied by Fourier Transform-Ion Cyclotron Resonance Mass Spectrometry

Supervisors: Prof. Peter B. O'Connor and Prof. Peter J. Sadler

Master of Chemistry (MChem)

University of Warwick October 2011 – June 2015

Thesis title: Studying and Developing Supercharging Effects and Techniques *via* High Resolution Tandem FT-ICR Mass Spectrometry

Supervisor: Prof. Peter B. O'Connor

Graduated Magna cum laude in 2015 (2.1 honours)

A Level

Berkhamsted School September 2009 – June 2011

Subject Taken: Chemistry (A2), Mathematics (A2), Further Mathematics (A2), Physics (A2), and Biology (AS)

Publications

Accepted papers

- Alessandro Del Grosso, Lavrentis-Dimitrios Galanopoulos, **Cookson K. C. Chiu**, Guy J. Clarkson, Peter B. O'Connor, Martin Wills, Strained alkynes derived from 2,2′-dihydroxy-1,1′-biaryls; synthesis and copper-free cycloaddition with azides, *Organic & Biomolecular Chemistry*, 2017, 15, 4517-4521, DOI: 10.1039/C7OB00991G
- Pingyu Zhang* and **Cookson K. C. Chiu***, Huaiyi Huang, Yuko P. Y. Lam, Abraha Habtemariam, Thomas Malcomson, Martin J. Paterson, Guy J. Clarkson, Peter B. O'Connor, Hui Chao, and Peter J. Sadler, Organoiridium Photosensitizers Induce Specific Oxidative Attack on Proteins within Cancer Cells, *Angewandte Chemie International Edition*, 2017, 56, 14898-14902, DOI: 10.1002/anie.201709082; *Angewandte Chemie International Edition* Inside Cover: DOI: 10.1002/anie.201710973
- Yuko P. Y. Lam, Christopher A. Wootton, Ian Hands-Portman, Juan Wei, **Cookson K. C. Chiu**, Isolda Romero-Canelon, Mark P. Barrow, Frederik Lermyte, and Peter B. O'Connor, Does Deamidation of Islet Amyloid Polypeptide Accelerate Amyloid Fibril Formation?, *Chemical Communications*, 2018, 54, 13853-13856, DOI:10.1039/C8CC06675B
- Christopher A. Wootton, Adam Millett, Andrea Lopez, **Cookson K. C. Chiu**, Mark P. Barrow, Guy J Clarkson, Peter J Sadler and Peter B. O'Connor. Structural analysis of peptides modified with organo-iridium complexes, opportunities from multi-mode fragmentation. *Analyst*, 2019, DOI: 10.1039/C8AN02094A
- "Determination of the Aggregate Binding Site of Amyloid Protofibrils using Electron Capture Dissociation Tandem MS" by Yuko P. Y. Lam, Christopher A. Wootton, Ian Hands-Portman, Juan Wei, **Cookson K. C. Chiu**, Isolda Romero-Canelon, Frederik Lermyte, Mark P. Barrow, and Peter B. O'Connor. Accepted to the Journal of the American Society for Mass Spectrometry

* authors contributed equally to the work published

Manuscripts in preparation

- "Superacid Supercharging using CH_5^+" Christopher A. Wootton and Haytham E. M. Hussein, Man Ying Wong, **Cookson K. C. Chiu**, Tomos E. Morgan, Mark P. Barrow, Peter B. O'Connor
- "The Next Dimension in Proteomics" by Yuko P. Y. Lam, Christopher A. Wootton, **Cookson K. C. Chiu**, Tomos E. Morgan, Remy Gavard, Maria A. van Agthoven, Meng Li, Mark P. Barrow, and Peter B. O'Connor.

Submitted papers

- "Metallodrug-Peptide Interactions Studied by Ultra-High Resolution Mass Spectrometry" by **Cookson K. C. Chiu**, Yuko P. Y. Lam, Christopher A. Wootton, Mark P. Barrow, Peter J. Sadler and Peter B. O'Connor. Submitted to Journal of the American Society for Mass Spectrometry
- "The Inhibition Mechanisms of Human Islet Amyloid Polypeptide Fibrillisation" by Yuko P. Y. Lam, **Cookson K. C. Chiu**, Christopher A. Wootton, Ji. I. Song, Meng Li, Ian Hands-Portman, Mark P. Barrow, and Peter B. O'Connor. Submitted to Chemical Communication

Application Notes and Patents

- Application note "Monitoring the Electrochemical Reduction of Disulphide Bonds in Proteins for Enhanced MS and MS/MS Capabilities". Christopher A. Wootton* and **Cookson K. C. Chiu***, Agnieszka Kraj, Thomas Lambert, Simon Lambert, and Peter B. O'Connor, in collaboration with ARC Sciences (UK) and Antec (Netherlands)
- Superacid supercharging ESI source, "IONISING MOLECULES AND ELECTROSPRAY IONISATION APPARATUS", Peter B. O'Connor, Christopher A. Wootton, Haytham Hussein, **Cookson K. C. Chiu**, Man-ying Wong, 2016, International Application Number: PCT/GB2017/050914

* authors contributed equally to the work published

Conference Presentations

Oral Presentations

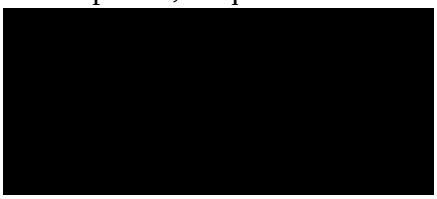
- International Mass Spectrometry Conference (IMSC 2018) – Florence, Italy

Poster Presentations

- European Fourier Transform Mass Spectrometry (EFTMS 2016) – Matera, Italy
- Bright Spark Symposium (2016) – Warwick, UK
- British Mass Spectrometry Annual Conference (BMSS, 2016) – Eastbourne, UK
- East Midlands Proteomic Workshop (EMPW 2016) – Nottingham, UK
- Analytical Research Forum (ARF 2017) – London, UK
- British Mass Spectrometry Annual Conference (BMSS, 2017) – Manchester, UK
- East Midlands Proteomic Workshop (EMPW 2017) – Nottingham, UK
- American Society of Mass Spectrometry Annual Conference (ASMS 2018) – San Diego, US
- Uppsala Conference on Electron Capture and Transfer Dissociation Mass Spectrometry (UppCon 2018) – Leeds, UK
- East Midlands Proteomic Workshop (EMPW 2018) – Lincoln, UK
- American Society of Mass Spectrometry Annual Conference (ASMS 2019) – Atlanta, US
- British Mass Spectrometry Annual Conference (BMSS, 2019) – Manchester, UK

Declaration

I hereby declare that except where specifically stated/reference's are made to other sources, the thesis entitled "Metallodrug Interactions with Proteins in Biological Systems Studied by Fourier Transform Ion Cyclotron Resonance Mass Spectrometry" is the original work of the named Author. It has been composed by myself and co-authors where stated and has not been submitted in whole or in part for any other degree, diploma, or qualification.



Cookson K. C. Chiu

July 2019

Abstract

The work presented herein focuses on the study of novel metallodrugs (transition metal-based drugs) and their interaction with different biomolecules in the form of peptides, proteins, and cellular components *via* the use of ultra-high resolution Fourier Transform-Ion Cyclotron Resonance Mass Spectrometry (FT-ICR MS).

A LC-MS/MS methodology with a specialised peak picking algorithm was developed for automatic detection of metallated species and information of metallodrug-binding was successfully elucidated. An alternative digestion method for bottom-up study was examined, however, standard tryptic digestion showed better and more predictable results. Further, the performance and efficiency of two distinct disulfide bond reduction methods: electrochemical reduction and radical-induced reduction were investigated. Results showed that both techniques can effectively reduce disulfide bonds of model proteins, opening the protein structure and enhancing fragmentation efficiency during MS/MS experiments to different controllable degrees; Better protein identification and characterisation was therefore afforded using these methods.

A range of octahedral iridium-(III)-based anticancer complexes were found to be capable of generating toxic singlet oxygen upon irradiation of visible blue light (463 nm), as evidenced by the production of a specific oxidised form of tryptophan, 3-hydroxykynurenine, on a model peptide. Proteomics studies of human lung cancer cells (A549) revealed the photoactivation of the iridium complex caused selective oxidative damage to specific histidine residues in key cellular proteins; Heat shock protein 70 kDa and aldose reductase.

The overall effects of an osmium-based anticancer drug against A2780 human ovarian cancer cells and an iridium-based antimicrobial complex against *Staphylococcus aureus* were studied using quantitative nLC-FT-ICR-MS/MS. The osmium-based anticancer complex affected the energy production pathway of the cancer cell, whereas the iridium-based antimicrobial complex was found to be capable of triggering a series of secondary responses from bacterial cells by inducing a vast amount of stress, including oxidative stress.

Table of Contents

List of Figures, Tables, and Schemes

Chapter 1: Introduction

Chapter 2: Metallodrug-Peptide Interactions Studied by Ultra-High Resolution Mass Spectrometry

Chapter 3: Reduction of Disulfide Bonds in Proteins for Enhanced MS and MS/MS Capabilities

Chapter 3: Supplementary Information

Chapter 4: Organoiridium Photosensitisers Induce Specific Oxidative Attack on Proteins within Cancer Cells

Chapter 4: Supplementary Information

Chapter 5: Metallodrug-Proteomics Pathway Analysis using nLC-FT-ICR MS

Chapter 5: Supplementary Information

Abbreviations

14-3-3 γ	14-3-3 Protein Gamma
14-3-3 ε	14-3-3 Protein Epsilon
14-3-3 ζ	14-3-3 Protein Zeta
$^1O_2$	Singlet Oxygen
1P	1-Photon
2P	2-Photon
3OH-kyn	3-hydroxykynurenine
5-ALA	5-aminolevulinic acid
AmmBic	Ammonium Bicarbonate
MeCN	Acetonitrile
AcpP	Acyl Carrier Protein
AE	α-Enolase
AHCY	Adenosylhomocysteinase
ahpC	Alkyl Hydroperoxide Reductase C
AR	Aldose Reductase
ATP	Adenosine Triphosphate
AUX	Auxiliary Electrode
BPGM	Bisphosphoglycerate Mutase
BSA	Bovine Serum Albumin
(C,N)	Phenylquinoline Ligand
C_A	Chain A of Human Insulin
CAD	Collisionally Activated Dissociation
C_B	Chain B of Human Insulin
CDDP	cis-diaminodichlorido-platinum/cisplatin
CEM	Chain Ejection Model
CF281	$[(Cp^{Xbiph})Ir(TolBig)Cl]Cl$
Chirp	Frequency-Sweep Excitation
CI	Chemical Ionisation
CID	Collisionally Induced Dissociation
HEWL	Chicken Egg White Lysozyme
CRM	Charged Residue Model
CRS	Charge Reduced Species
DAP	L,L-2,6-Diaminoheptanedioate
dapE	Succinyl-Diaminopimelate Desuccinylase
DAVID	Database for Annotation, Visualisation and Integrated Discovery
DC	Direct Current
DHAP	Dihydroxyacetone
DNA	Deoxyribose Nucleic Acid
DTT	Dithiothreitol
EC	Electrochemistry/Electrochemical

ECD	Electron Capture Dissociation
EDD	Electron Detachment Dissociation
EI	Electron Impact / Electron Ionisation
EIC	Extracted Ion Chromatogram
EID	Electron Induced Dissociation
ESI	Electrospray Ionisation
ETD	Electron Transfer Dissociation
ExD	Electron-based Dissociation Techniques
FBA	Fructose-Biphosphate Aldolase
FDA	Food and Drug Administration
FDR	False Discovery Rate
FT-ICR MS	Fourier Transform-Ion Cyclotron Resonance Mass Spectrometry
FWHM	Full Width Half Maximum
FY12	$[(\eta^6\text{-bip})Os(azpy\text{-}NMe_2)I]PF_6$
G3P	Glyceraldehyde 3-Phosphate
GC	Gas Chromatography
Glycerate-1,3P2	1,3-bisphosphoglycerate
Glycerate-2,3P2	2,3-Bisphosphoglycerate
Glycerate-2P	2-Phosphoglycerate
Glycerate-3P	3-Phosphoglycerate
GPI	Glucose-6-Phosphate Isomerase
hECD	Hot Electron Capture Dissociation
HILIC	Hydrophilic Interaction Liquid Chromatography
hLZ	Human Lysozyme
Hsp 70	Heat Shock Protein 70 kDa
HSP90AA1	Heat Shock Protein HSP 90-Alpha
HSP90B1	Endoplasmin
Hz	Hertz
IAA	Iodoacetamide
IS	Internal Standard
IC_{50}	Half-Maximal Inhibitory Concentration
ICR	Ion Cyclotron Resonance
IEM	Ion Ejection Model
IR-ECD	Infrared-Activated Electron Capture Dissociation
IRMPD	Infra-red Multi-Photon Dissociation
IsaA	Immunodominant Staphylococcal Antigen A
iTRAQ	Isobaric Tag for Relative and Absolute Quantification
KE	Kinetic Energy
KEGG	Kyoto Encyclopedia of Genes and Genomes Database
L3BBS	$[Lys^3]$-Bombesin
LC	Liquid Chromatography
LED	Light-Emitting Diode
L-LDH	Lactate Dehydrogenase

LP Hg	Low Pressure Mercury
m/z	Mass-to-charge ratio
MALDI	Matrix Assisted Laser Desorption Ionisation
MAT2A	S-Adenosylmethionine Synthase Isoform
MCTSs	Multicellular Tumour Spheroids
MDR	Multiple Drug Resistance
MIC	Minimum Inhibitory Concentration
MoA	Mechanism of Action
mRNA	Translation Messenger Ribonucleic Acid
MS	Mass Spectrometry
MS/MS	Tandem Mass Spectrometry
NADPH	Nicotinamide Adenine Dinucleotide Phosphate (reduced form)
nCI	Negative Chemical Ionisation
NEM	N-ethylmaleimide
nESI	Nano Electrospray Ionisation
NHS	N-hydroxysuccinimide
nLC	Nano Flow Liquid Chromatography
(O,O)	Diketonate Ligand
Os1	Os1-Cl after loss of the chlorido ligand and counterion
Os1-Cl	$[(\eta^6\text{-bip})Os(en)Cl]PF_6$
PACT	Photoactivatable Chemotherapy
PCK	Phosphoenolpyruvate Carboxykinase
PDT	Photodynamic Therapy
PEP	Phosphoenol Pyruvate
pflB	Formate acetyltransferase
PGK	Phosphoglycerate Kinase
PK	Pyruvate kinase
ppb	Parts-per-billion
ppm	Parts-per-million
PTM	Post-Translational Modification
PZ6	$[Ir(phenylquinoline)_2(diketonate)_2]$
q	Charge of Ion
Qh	Quadrupole-Hexapole
REF	Reference Electrode
RF	Radio Frequency
RNA	Ribonucleic Acid
RNAase A	Ribonucleuase A
ROS	Reactive Oxygen Species
RP	Reversed Phase
rpoE	Ribonucleic Acid Polymerase
(S,S)	Dithione Ligand
S. aureus	*Staphylococcus aureus*
S/N	Signal-to-noise ratio

SAX	Strong Anion Exchange
sbi	Immunoglobulin-Binding Protein sbi
SCX	Strong Cation Exchange
SDAP	N-Succinyl-L,L-Diaminopimelic Acid
SDS	Sodium Dodecyl Sulphate
SILAC	Stable Isotope Amino Acids into the Cell Culture
SNAP	Sophisticated Numerical Annotation Procedure
SNAP-LC	SNAP-Liquid Chromatography Algorithm
SWIFT	Stored Waveform Inverse Fourier Transform
TCEP	Tris(2-carboxyethyl)phosphine
TMT	Tandem Mass Tag
ToF	Time-of-Flight
TPI 1	Triosephosphate Isomerase
tpx	Thiol Peroxidase
tRNA	Translation Ribonucleic Acid
V	DC Voltage
UV	Ultraviolet
UVPD	Ultraviolet Photodissociation
f	RF frequency
WE	Working Electrode
α-D-Glucose-6P/β-D-Glucose-6P	Glucose-6-Phosphate
β-D-Fructose-6P	Fructose-6-Phospate

Chapter 1 Introduction

1.1 Introduction of Mass Spectrometry

Mass spectrometry (MS) is an important analytical technique to study different molecules/species based on their masses. For molecules/species to be detected, they need to be charged, and therefore, MS measures their mass-to-charge ratios (*m/z*).

The first study based on charged ions of a molecule was achieved by Sir J. J. Thompson on O_2, N_2, CO, CO_2, and $COCl_2$ in 1912. MS since then has developed rapidly for the last 100 years and has become one of the most employed analytical techniques to study a wide range of research areas, including proteomics,[1] metabolomics,[2] and petroleomics.[3] The mass spectrometer consists of three major components, which are ionisation source, mass analyser, and detector. An introduction of ionisation techniques, mass analysers and fragmentation methods will be presented in this chapter.

1.1.1 Ionisation Techniques in Mass Spectrometry

As only charged gas state molecules can be detected in a mass spectrometer, it is important to ionise the molecules prior to detection. In general, ionisation techniques can be grouped into two categories: hard ionisation and soft ionisation. Hard ionisation refers to the ionisation processes being energetic, and fragments from the molecular ions are generated during the process.[4] Fragments generated during the hard ionisation process are useful and can be used for structural elucidation of the analytes. On the other hand, due to the generation of multiple fragments per molecular ion, hard ionisation is not the best option if a mixture of analytes is involved as it can introduce more challenges to the data analysis. Soft ionisation, on the other hand, is milder as an ionisation technique compared to hard ionisation. Mostly molecular ions are preserved during the ionisation process, avoiding the complication of data analysis if complex mixture of analytes is involved, and enabling the detection of more labile/fragile species.

1.1.1.1 Electron Impact/ Ionisation

Electron impact/ ionisation (EI) is one of the hard ionisation techniques in MS and was introduced by Sir A. J. Dempster in 1918.[5] In EI ionisation, electrons are generated from a heated filament which has electric current running through it. Electrons are accelerated to 70 eV (1 eV= 96.485 kJmol^{-1}) and attracted towards an anode in an ionisation chamber (see Figure 1.1). Vaporised/gaseous molecules are introduced from the inlet which is placed orthogonally to the heated filament and collide with the high energy electron beam to displace a bound electron from the analyte species, forming a radical cation ($M^{+\bullet}$). The overall EI ionisation process can be represented by the equation below:

$$M + e^-_{fast} \rightarrow M^{+\bullet} + 2e^- \qquad \text{[Eq. 1]}$$

M represents the gaseous molecule, e^-_{fast} represents the high energy electron, and $M^{+\bullet}$ represents the radical cation.

The radical cations generated by the EI ionisation process are usually unstable and can undergo a series of dissociation reactions and generate more stabilised fragment ions. The extensive fragment ions generated can be used for structural elucidation of the analytes. Gas chromatography (GC) is commonly used to couple with EI ionisation to separate a mixture of analytes based on the difference in boiling points of the components in the mixture. EI ionisation is generally employed to volatile small molecules, thus suitable for clinical, pesticides, food/drug analysis, and other applications.[6,7] The low volatility of most biomolecules is also a limiting factor in EI ionisation of biomolecules.

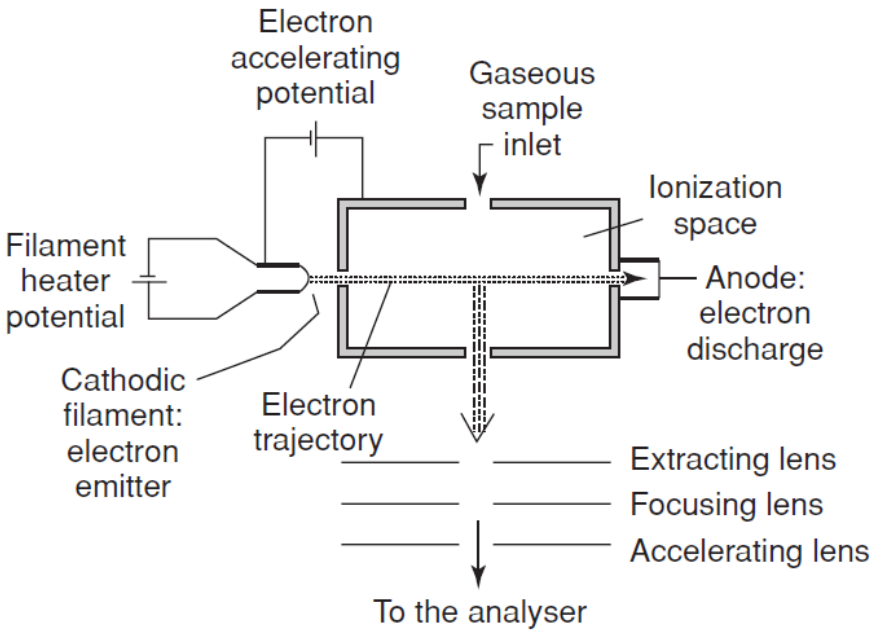


Figure 1.1 Schematic representation of an electron ionisation source (Reproduced from De Hoffmann *et al.* 2007 with permission).[4]

1.1.1.2 Chemical Ionisation

Chemical ionisation (CI) was introduced by Talrose in 1952[8] and further developed extensively by Munson and Field.[9,10] Unlike EI ionisation, CI was designed to be a proton-transfer based ionisation source, instead of an electron based method. Highly reactive protonated gas/reagent gas is used to promote proton transfer from the reagent gas to the analyte species thus creating protonated ions as a result compounds with extremely low proton affinity are desired, i.e. extremely reactive Brønsted acids, meaning the capability of donating their extra proton effectively. The formation of gas-phase proton-transfer reagent gas is usually achieved through a glow discharge source/chamber. A high energy spark converts methane (CH_4) into an extremely reactive Lewis acid, sometimes referred as superacid, CH_5^+, this process is represented by the equations below:

Step 1: CI reagent gas – methane (CH_4)

$$CH_4 + e^- \rightarrow {CH_4}^{+\bullet} + 2e^- \quad \text{[Eq. 2]}$$

$${CH_4}^{+\bullet} + CH_4 \rightarrow {CH_5}^+ + {CH_3}^\bullet \quad \text{[Eq. 3]}$$

Step 2: Proton transfer to analyte

$$M + {CH_5}^+ \rightarrow MH^+ + CH_4 \quad \text{[Eq. 4]}$$

Compared to EI ionisation, CI is a milder/softer ionisation technique, meaning CI causing less/little to no fragmentation to analyte species, generating predominantly MH^+ ions. However, CI only produce singly charged ions and is a gas phase reaction, imparting volatility requirements, limited its use to smaller analyte species only and preventing the study of larger biomolecules, such as peptides and proteins.

1.1.1.3 Electrospray Ionisation

Electrospray ionisation (ESI) was first introduced by John Fenn in 1985 who was later awarded a Nobel Prize in Chemistry in 2002.

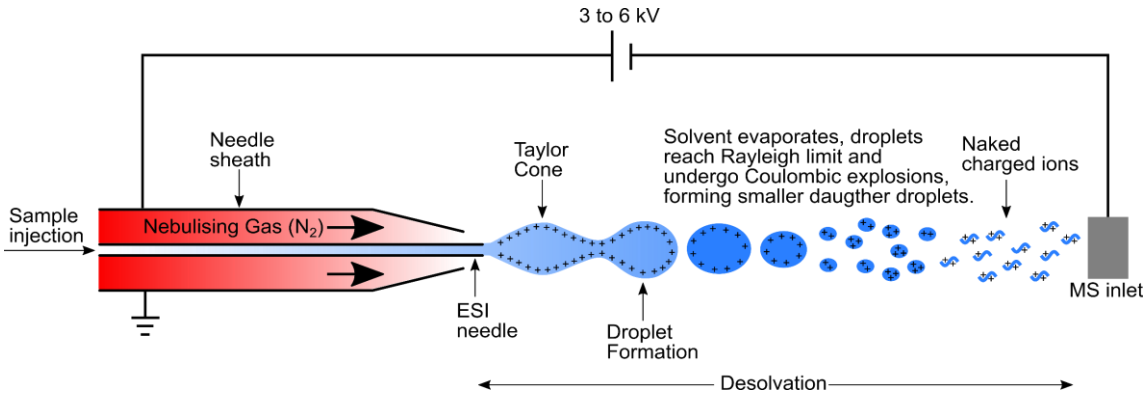


Figure 1.2 Schematic representation of an ESI ionisation source. Droplets undergo desolvation, reach their Rayleigh limits, explode into smaller daughter droplets, eventually form naked charged ions.

The analyte solution is first exposed to a high potential difference (typically 3-6 kV) between the solution and the mass spectrometer inlet. The charged analyte solution

is then passed through an ESI needle, mixed with an inert gas (often nitrogen) and form a Taylor cone, which then leads to the formation of highly charged droplets. The charged droplets within the plume will be attracted towards the mass spectrometer inlet. As the charged droplets move towards to mass spectrometer inlet under the electric field, desolvation of the droplet occurs. When the solvent surrounding the charges in the droplet evaporates, the number of charges in the droplet stays approximately constant, and the droplets eventually approach the Rayleigh limit of instability.[11] When a droplet approaching its Rayleigh limit (see Eq. 5 below), it will fission into two or more species (daughter droplets), average of 13 daughter droplets per coulombic explosion was observed in some cases.[12] The daughter droplets will continue to move towards the mass spectrometer inlet under the electric field like before and desolvation continues, until reaches their Rayleigh limits, and fission happens again. The process repeats until eventually producing desolvated ions. The desolvated ions can be manipulated and detected by the mass spectrometer and analysed.

$$z_r e = \sqrt{8\pi(\epsilon\gamma r^3)} \qquad \text{[Eq. 5]}$$

where z_r represents the charge of electrospray droplet at Rayleigh limit, r is the radius of the electrospray droplet at Rayleigh limit, ε represents permittivity of free space, and γ is the surface tension of the solute.

ESI is a very soft ionisation technique, meaning the ionisation process only produces molecular ions of the analytes, $[M+nH]^{n+}$ for positive ion mode or $[M-nH]^{n-}$ for negative ion mode, without inducing modifications or fragmentation to the analytes. The ability of producing multiply charged analyte species by ESI allows the feasibility of analysing large biomolecules, *e.g.* large proteins or deoxyribose nucleic acid (DNA), by increasing the charges on the molecules, lowering the resulting *m/z* ratio into a more achievable mass range for most mass spectrometers. However, the mechanism of analyte charging, generating naked charged ion from solvated-analyte droplets is still under debate despite the fast development of MS in the past century. Several models have been proposed; the ion ejection model (IEM), the charged residue model (CRM), and the chain ejection model (CEM), are the most widely accepted models (Figure 1.3).[13]

Konermann *et al.* suggested that IEM is mainly applicable for small molecules, usually getting charged *via* proton/salt ions already present in the solution. CRM is mainly employed by large molecules such as folded proteins and protein complexes which require a lot of energy to overcome the energy barrier for IEM to occur. CEM is mainly applicable for unfolded proteins (disordered/extended chain structure), such as denatured biomolecules and polymers.

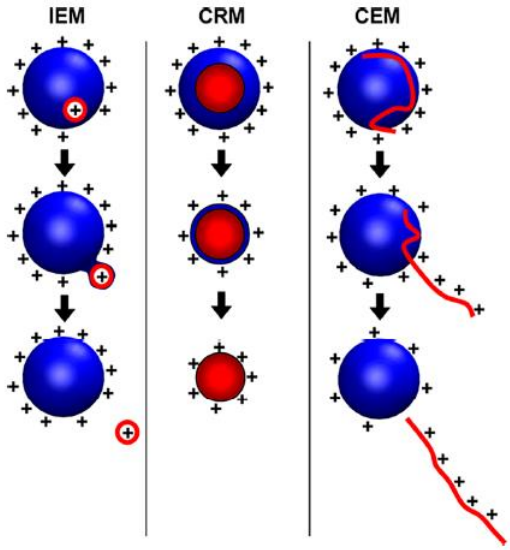


Figure 1.3 Summary of the three most widely accepted models for ESI ion formation. From left to right, the Ion Ejection Model, the Charge Residue Model, and the Chain Ejection Model. (Reproduced from Konermann *et al.* 2013 with permission).[13]

1.1.1.4 Nano Electrospray Ionisation

Nano electrospray ionisation (nESI) has been developed to decrease sample consumption, while retaining a stable, reliable ion signal, i.e. higher sensitivity. The orifice diameter of the nESI emitter (~0.5-5 μm) is much smaller than that of a standard ESI emitter (~10-100 μm).[14] Therefore, the initial droplet size forming from nESI is inherently smaller than those forming from ESI. Unlike ESI, nESI does not require a nebulising inert gas to achieve an effective ESI plume. Due to the advantage of smaller initial droplet size, nESI emitters are generally placed a lot closer to the MS inlet compared to ESI as the process does not require as much time and space for the droplet

to expand and produce naked charged ions. nESI typically uses much lower voltages to achieve effective ionisation (~500-1500 V) compared to ESI, which usually requires 3000-6000 V. As initial droplet size is smaller in nESI, lower source temperature can also be used for effective desolvation, providing a more suitable option for native MS of biomolecules, where excess heat from desolvation may cause denaturation. nESI emitters are generally pulled borosilicate glass capillaries, which can be easily made using a heated capillary puller. The shape of the capillaries (orifice diameter, tapering etc) can be optimised using different parameters, such as temperature of the heated filament, pulling velocity etc. nESI emitters generally require optimisation depending on the sample nature or different solvent system. A schematic representation of nESI is in Figure 1.4.

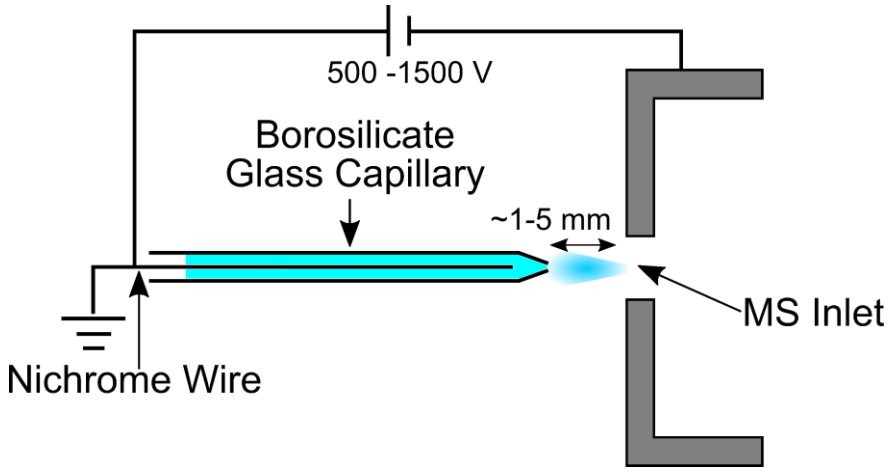
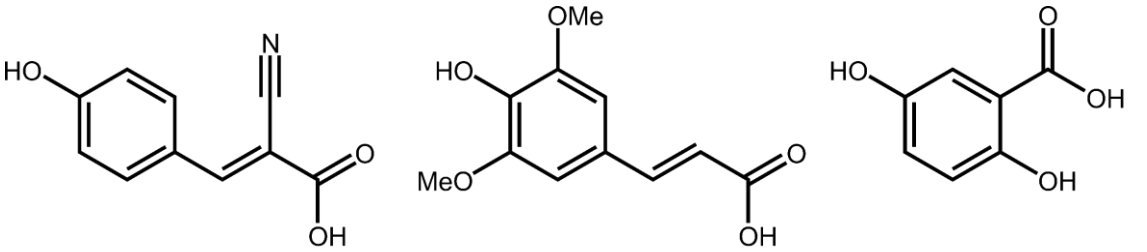


Figure 1.4 Schematic representation of a nESI experiment setup.

1.1.1.5 Matrix Assisted Laser Desorption Ionisation

Matrix assisted laser desorption ionisation (MALDI) was invented by Karas and Hillenkamp in 1985.[15,16] Tanaka *et al.* later showed species over 100,000 *m/z* can be analysed by MALDI MS.[17] MALDI uses a solid sample, which is ionised inside a chamber under vacuum. Desorption and ionisation are the two most important steps in MALDI. To achieve effective desorption in MALDI, matrices are required. Matrices are composed of aromatic rings and a functional group that could provide a proton for the analyte (see examples of matrices below in Figure 1.5).



α-cyano-4-hydroxycinnamic acid (CHCA) | sinapinic acid (SA) | 2,5-dihydroxybenzoic acid (DHB)

Figure 1.5 Examples for 3 common matrices used in MALDI-MS.

Analytes are first mixed effectively with matrix solution. Then the mixed sample is spotted onto a stainless steel MALDI target plate and allowed to dry. When the sample is irradiated with laser, aromatic groups on the matrix molecules in the mixed sample will absorb a large amount of energy and deprotonate. The released protons hopefully transfer to the neutral analyte molecules and ionise them. The analyte ions can then be manipulated and detected with a mass analyser. Like ESI and nESI, MALDI is also a soft ionisation technique, meaning the ionisation process does not induce excessive fragmentation. However, the excess laser energy can still lead to fragmentation of the analyte molecular ions in certain cases. Usually only singly charged ions are generated from MALDI, which restricts the optimal detection from many instruments. Low shot-to-shot reproducibility is also a main disadvantage compared to ESI and nESI. A schematic representation of MALDI is shown below. (Figure 1.6).

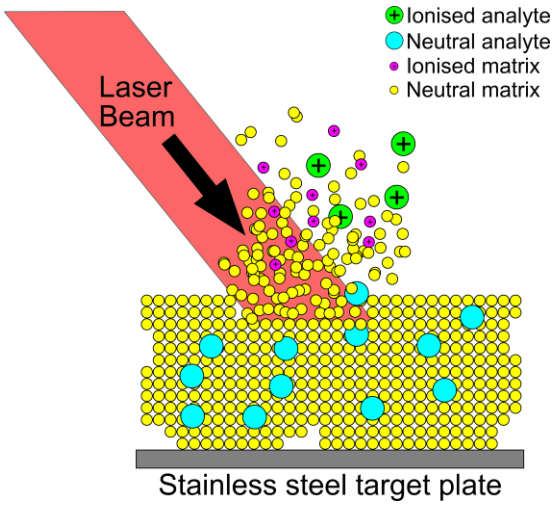


Figure 1.6 Schematic representation of MALDI ionisation process. N_2 laser activates and ionises matrix molecules into ions, which then undergo proton transfer with nearby analyte molecules, generating detectable analyte ions.

1.1.2 Mass Analysers

After the generation of ions, they can be manipulated and transmitted down to the mass analyser using electric potential differences. Ions are separated according to their corresponding *m/z*, and finally getting detected by the mass analyser. There are many types of mass analysers, for brevity the following three are discussed, quadrupole, time-of-flight (ToF), and Fourier transform-ion cyclotron resonance (FT-ICR) MS. Mass accuracy and resolving power are often measured to evaluate the performance of a mass analyser.

Mass accuracy measures the deviation of the measured *m/z* value of a certain analyte species from its theoretically calculated *m/z*. Mass accuracy of a mass analyser is expressed in parts-per-million (ppm), and can be represented by the equation below:

$$Mass\ Accuracy\ (ppm) = \frac{(Measured\ m/z - Theoretical\ m/z)}{Theoretical\ m/z} \times 1{,}000{,}000 \qquad \text{[Eq. 6]}$$

Resolving power is a measure of how effectively a mass spectrometer can separate mass spectral peaks, the higher the resolving power, the more effective the

instrument is at resolving closely spaced peaks. The resolving power of a mass spectrometer can be calculated using the equation below:

$$Resolution\ (R) = \frac{m}{\Delta m} \qquad \text{[Eq. 7]}$$

m represents the *m/z* of the peak in question, and Δm represents the width of the peak at half of its own height, i.e. its Full Width Half Maximum (FWHM).

1.1.2.1 Quadrupoles Mass Analyser

The quadrupole analyser was introduced by Paul and Steinwedel in 1953.[18] A quadrupole is generally made of 4 circular rods which are placed parallel to each other. Ions travel along the z-axis and oscillate in the x and y planes. The electric field generated by the direct current (DC), voltage (V), and radio frequency (RF) amplitude (V) on the 4 rods dictates the oscillation of the ions. The 4 rods are electronically connected oppositely in pairs and the potential on the different pairs of rods are set to be completely out of phase, i.e. 180^{o} and rapidly oscillate between the pairs of rods (see Figure 1.7).[4]

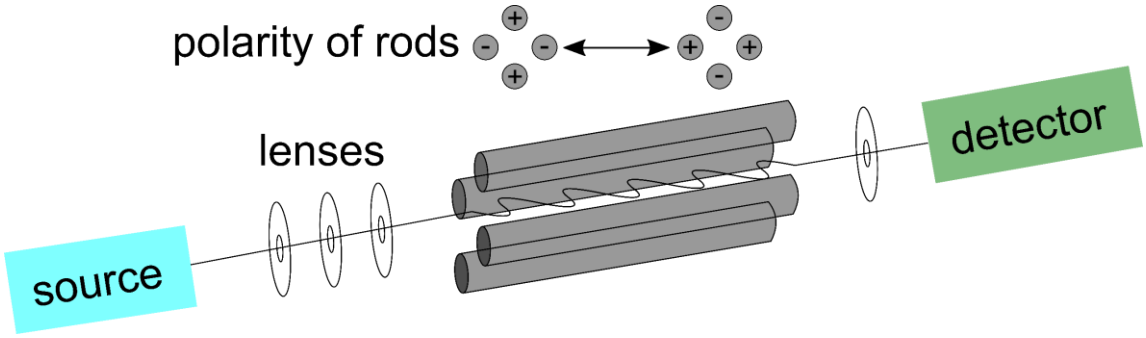


Figure 1.7 Schematic representation of a quadrupole, 4 parallel rods are paired up electronically with the rod opposite to each other.

Ions are accelerated down the electric gradient before entering the quadrupole, then retain the same velocity in the z axis during transmission. The trajectory stability of ions can be represented by the two equations below (Eq. 8 and Eq. 9).[19]

$$a = \frac{8zV}{mr_0{}^2\omega^2} \qquad \text{[Eq. 8]}$$

$$q = \frac{4zA}{mr_0{}^2\omega^2} \qquad \text{[Eq. 9]}$$

where a and q are directly proportional to V and A respectively, z represents charge of the ion, m represents mass of the ion, r_0 is the field radius, and ω is the angular frequency.

Only ions with stability at certain V and A values can be transmitted and detected, while all other unstable ions will be discharged, *e.g.* onto the rods, getting neutralised, i.e. cannot be detected. The principle of ion stability in a quadrupole can be represented *via* the Mathieu stability diagram (see Figure 1.8).

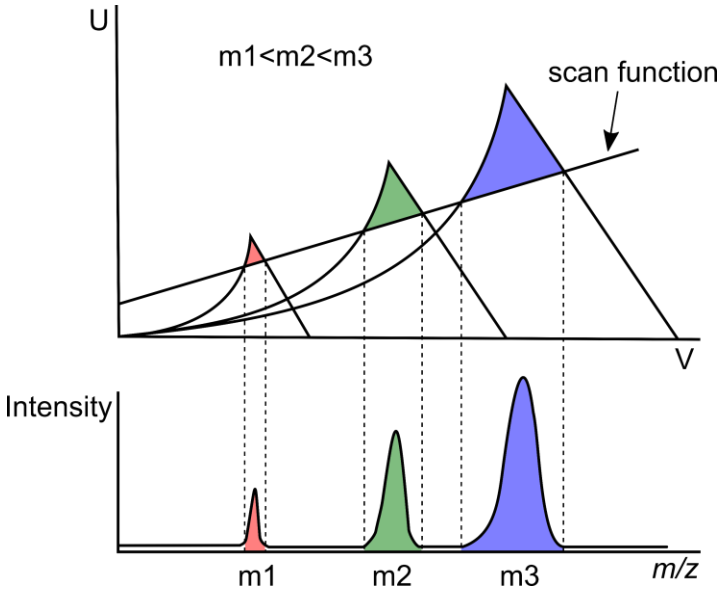


Figure 1.8 Mathieu Stability Diagram. The intersection of the scan function with the stability diagram is directly related to the width of the spectral peak. The areas under the curve is corresponding to the spectral peak area. By altering the DC voltages (V) and RF frequency (f) applied to the quadrupole, the scan function changes, thus the regions of stability of each species changes.

The resolving power of a quadrupole is approximately 2000 at 1000 *m/z*.[4] Quadrupole are well known for their robustness and reliability, therefore it has been employed in many different research fields.[20] Quadrupoles can also be coupled to other

mass analysers, acting as an ion guide, mass filter, or a collision cell, which will be discussed further in section 1.1.3.

1.1.2.2 Time-of-flight Mass Analyser

The principle of the time-of-flight instrument was first proposed by Stephens back in 1946.[21] In ToF-MS, ions are accelerated through a flight tube by an electric field, induced by the potential difference between the electrodes and the extraction grid. In theory, all ions gain the same amount of kinetic energy, however different ions have different masses, therefore travel with different velocity and reach the detector at different times.

Kinetic energy gained by ions:

$$Kinetic\ energy\ (KE) = \frac{1}{2}mv^2 = qV \qquad \text{[Eq. 10]}$$

Rearranging equation 10:

$$v = \sqrt{\frac{2qV}{m}} \qquad \text{[Eq. 11]}$$

where m is the mass of the ion, v is the velocity of the ion, q = ze is the charge of the ion, and V is the electric potential generated from the potential difference between the electrodes and the extraction grid.

From equation 11, it shows that the ion velocity gained from the electric field is inversely proportional to the mass of the ion, i.e. heavier ions travel slower and vice versa. In theory all ions get the same amount of kinetic energy, however the kinetic energy gained by each ion is slightly different to each other in reality, leading to spreading of ions drift time, even ions with the same *m/z*, causing peak broadening and low peak resolution. To overcome this, multiple solutions have been developed. The reflectron is a common technique employed to correct the difference in kinetic energy between ions. The idea of reflectron in ToF was first introduced by Mamyrin *et al.* in

1973.[22] The reflectron acts like a mirror which reflects the ions travelling from the ionisation source to the detector (seen Figure 1.9).

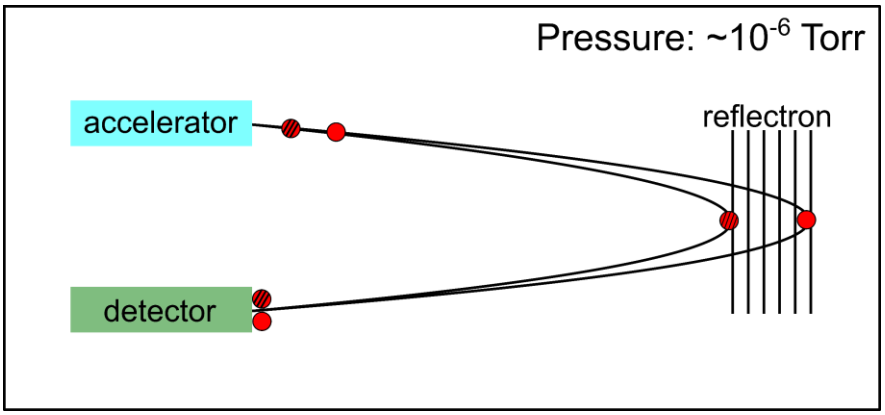


Figure 1.9 Schematic representation of a reflectron ToF. Ions with higher kinetic energy (solid sphere) travel faster and penetrate deeper into the reflectron while ions with lower kinetic energy (striped sphere) travel slower and penetrate not as deep into the reflectron compared to ions with higher kinetic energy. The time difference of the ions spending in the reflectron corrects the spreading of kinetic energy, resulting in ions with the same *m/z* arriving the detector as the same time.

Ions with higher kinetic energy travel faster and penetrate deeper into the reflectron compared to ions with less kinetic energy, meaning ions with higher kinetic energy will spend a longer time in the reflectron than that of ions with lower kinetic energy. Therefore, ions with the same *m/z*, even with different initial kinetic energy after leaving the source can arrive the detector at the same time. With the application of reflectron in ToF-MS, the mass accuracy can be enhanced significantly from 200 ppm to 10 ppm (depending on the calibration), and resolving power at 1000 *m/z* can be improved greatly from 5,000 to 20,000 or higher, resolving power of 100,000 can be achieved with some commercial setups now.[4]

1.1.2.3 Fourier Transform-Ion Cyclotron Resonance Mass Analyser

Fourier transform-ion cyclotron resonance (FT-ICR) mass analyser was first developed by Marshall and Comisarow in 1974.[23,24] FT-ICR is well-known for its high

mass accuracy and resolving power, therefore it has been widely applied to many scientific researches.[3,25–27] The major principles of FT-ICR will be discussed below.

1.1.2.3.1 Cyclotron Motion

Cyclotron motion describes the ion motion in the presence of a magnetic field. The ICR cell is placed in the centre of a superconducting magnet which creates a homogenous magnetic field across the ICR cell. When the ions are transmitted into the ICR cell, with the motion perpendicular to the existing magnetic field, the ions experience a force called the "Lorentz force" (see Figure 1.10). The Lorentz force can be expressed by the equation below:

$$F = qv \times B \qquad \text{[Eq. 12]}$$

where F is the calculated Lorentz force experienced by the ion of interest, q is the charge of the ion, v is the velocity of the ion, and B is the magnetic field strength created by the superconducting magnet.

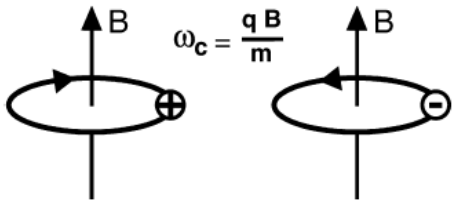


Figure 1.10 Cyclotron motion of a positively charged ion (left) and a negatively charged ion (right) in the presence of magnetic field (B). (Reproduced from Marshall *et al.* 2002 with permission).[28]

With the effect from the Lorentz force experienced by the ions under the magnetic field, the ions orbit around the centre axis of the magnetic field (z-axis), and this motion is referred as the "cyclotron motion". The theoretical cyclotron frequency of an ion orbiting the centre of the magnetic field can be represented with the equation below:

$$\omega_c = \frac{qB}{2\pi m} \qquad \text{[Eq. 13]}$$

where ω_c is the cyclotron frequency of the ion in Hertz (Hz), q is the charge of the ion, B is the magnetic field strength, and m is the mass of the ion.

In equation 13, it is shown that the cyclotron frequency of an ion is inversely proportional to the *m/z* of the ion, meaning higher *m/z* ions orbits the centre axis of the magnetic field at a lower frequency and vice versa. Based on the difference in cyclotron frequencies of ions, ions with different *m/z* can be separated. In theory, if cyclotron motion is the only contributing factor of the ion movement, the cyclotron frequency of an ion can be accurately measured, then by applying a calibration function, the measured cyclotron frequency can be expressed in *m/z* value. However, to measure the cyclotron frequencies of the ions in the ICR cell, they must be trapped in the cell.

1.1.2.3.2 Trapping Motion

An ICR cell consists of four main components; trapping electrodes, excitation electrodes, detection electrodes, and an amplifier to amplify small magnitude signals up to a transmittable level (see Figure 1.11).[29]

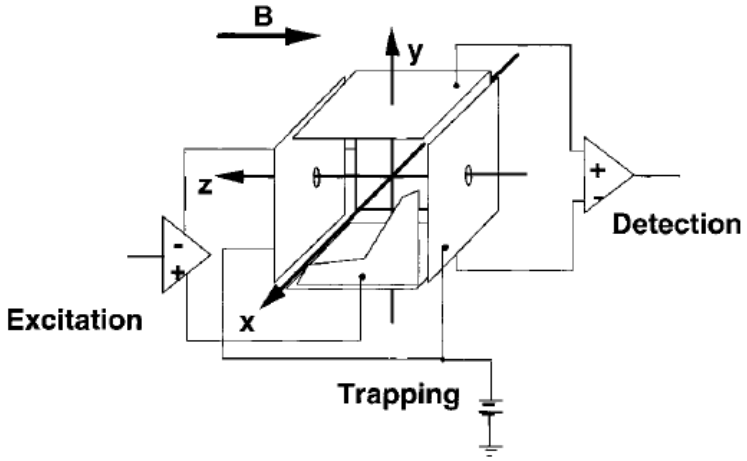


Figure 1.11 Schematic representation of a cubic ICR cell, showing the four major components; trapping electrodes, excitation electrodes, detection electrodes and an amplifier (Reproduced from Marshall *et al.* 1998 with permission).[29]

Without a trapping motion, ions will escape the FT-ICR cell along the z-axis. To avoid the ions from escaping the cell, a low electric potential of equal polarity can be applied to the two trapping electrodes of the cell, creating a "potential wall". When the energies of the ions are lower than the trapping voltages, the ions will be confined axially within the limits of the ICR cell, oscillating back and forth between the electrodes. The trapping motion frequencies of ions are around several kHz and can be represented by the equation below:[29]

$$\omega_z = \sqrt{\frac{2qV_{trap}\alpha}{ma^2}} \qquad \text{[Eq. 14]}$$

where ω_z is the trapping oscillating frequency induced by the trapping motion, q is the charge of the ion, V_{trap} is the trapping potentials applied to the trapping electrodes, m is the mass of the ion, α is geometry factor of the ICR cell, and a is the plate-to-plate difference of the ICR cell.

1.1.2.3.3 Magnetron Motion

Introducing an electric potential to the trapping plates can effectively confine the ions in the ICR cell, however, this also induces an electric field between the trapping plates. The combination of the electric and magnetic forces on the ions produces a motion called the "magnetron motion". The magnetron motion acts along the same axis as the cyclotron motion but at a much lower frequency (<100 Hz), and can be represented by equation 15 shown below:

$$\omega_m = \frac{V_{trap}\alpha}{\pi a^2 B} \qquad \text{[Eq. 15]}$$

where ω_m is the magnetron frequency in Hz, V_{trap} is the trapping potentials applied to the trapping plates, α is the geometry factor of the ICR cell, a is the distance between the two trapping plates, and B is the magnetic field strength.

Since the cyclotron motion and magnetron motion are acting along the same axis, the frequency of the ion measured by an ICR cell, known as the "reduced cyclotron frequency" can be represented with the equation below:

$$\omega_{measured} = \omega_c - \omega_m \qquad \text{[Eq. 16]}$$

The summarised ion motions (cyclotron motion, trapping motion, and magnetron motion) in a penning trap is shown below (see Figure 1.12)[29]

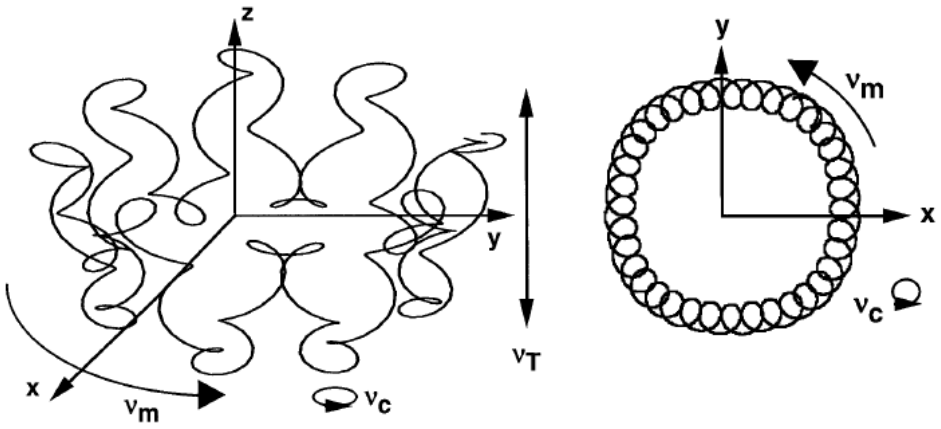


Figure 1.12 Overall ion trajectory with the contribution of the cyclotron motion (ν_C), the trapping motion (ν_T), and the magnetron motion (ν_M) in a penning trap (Reproduced from Marshall *et al.* 1998 with permission).[29]

Motion of ions discussed above assumes an ideal vacuum, which is unachievable. Damping of the three principle ion motions can be caused by ion-ion interactions and ion-background gas collision, with the cyclotron motion being particularly susceptible by these effects. The damping of cyclotron motion is represented below in Figure 1.13, showing with collision of excessive background gas can lead to loss of ions from the ICR cell. Therefore, the use of ultra-high vacuum conditions is crucial for measurements in FT-ICR MS.

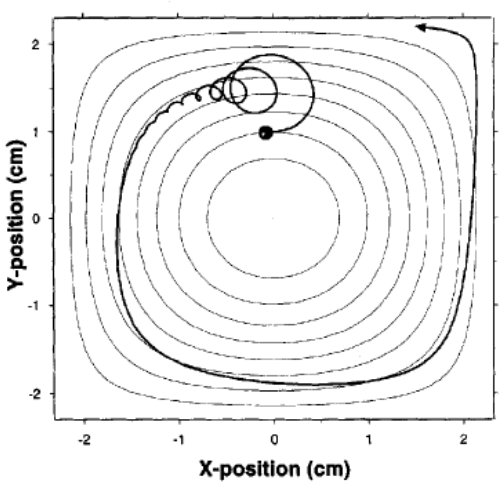


Figure 1.13 Representation of damping of an ion's cyclotron motion with excessive background gas (pressure= 0.01 Torr). The ion is quickly lost from the ICR cell. (Reproduced from Amster *et al.* 1996 with permission).[30]

1.1.2.3.4 Excitation and Detection in FT-ICR Cell

When ions enter an ICR cell, the cyclotron radii of the ions are generally too small to generate a detectable signal and are incoherent between each other, i.e. ions with the same frequency are not travelling together thus net image current is zero. A radio frequency (RF) can be applied to the two excitation electrodes (see Figure 1.11) to excite the ions. The RF pulse applied on the excitation electrodes, thus resonant excitation, will transfer energy to the ions trapped in the ICR cell. From equations 13-16, it is shown that the cyclotron frequency of an ion is independent of its kinetic energy, therefore added energy from the RF pulse will allow the ions to travel faster and be excited to achieve a larger cyclotron radius, while the cyclotron frequency remain the same around the ICR cell. The excitation event in an ICR cell is summarised in Figure 1.14 below:

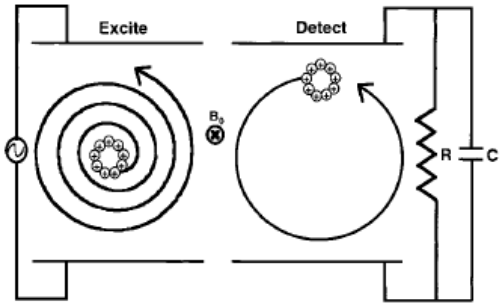


Figure 1.14 Representation of an ion cloud being excited by a RF pulse, showing acceleration with a spiral trajectory (left) into a large, detectable orbit (right) (Reproduced from Marshall *et al.* 1998 with permission).[29]

To excite a wide range of frequencies, thus a wide *m/z* range, approaches of different excitation profiles have been proposed. Frequency-sweep excitation (chirp) and stored waveform inverse Fourier transform (SWIFT) are the two most common approaches applied in FT-ICR MS. The chirp was first introduced by Marshall in 1974,[24] which excites ions over a broad frequency sweep with a low excitation voltage.[31] The chirp provides a wide frequency excitation in a straight forward fashion. However, it also exhibits excitation frequency profiles which are not completely uniform/flat in the region of interests, especially at the beginning and end of each pulse (horns) (see Figure 1.15c). Over-excitation of ions in these regions influences the observed relative abundance of the ions with those frequencies. To overcome the non-uniform excitation profile by the chirp, a flat consistent excitation profile would be ideal. SWIFT was proposed by Marshall later in 1985 to overcome the short-comings of chirp excitations.[32] Instead of applying a frequency sweep to a wide frequency range, SWIFT predicts the perfect excitation profile (rectangular excitation) then uses an inverse Fourier transform to calculate the time domain signal required obtain a perfect rectangular excitation profile (see Figure 1.15d). SWIFT can further be applied to isolate the *m/z* of interest in an ICR cell by over exciting the ions with undesired *m/z* by over-exciting them into larger orbits towards the electrodes in the ICR cells, getting neutralised and lost, while only the ions of interest are retained in the ICR cell (see Figure 1.15e).[29]

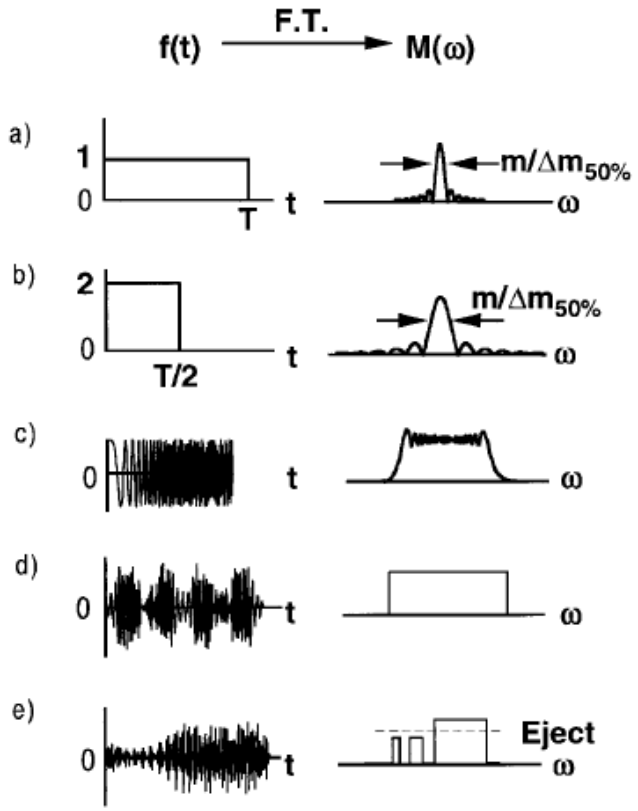


Figure 1.15 Different excitation pulses used to excite ions into a larger orbit in an ICR cell and enable detection. a) and b) show rectangular pulses with different magnitudes and durations, c) represents a standard frequency-sweep excitation (chirp) used in most FT-ICR MS instruments, d) Stored Waveform Inverse Fourier Transform (SWIFT), and e) SWIFT isolation/ejection excitation profile (Reproduced from Marshall *et al.* 1998 with permission).[29]

Once ions are excited to a large enough orbit, and thus become detectable inside the ICR cell, the RF pulse applied to the excitation electrodes is turned off. Then the ions continue to orbit around the ICR cell at the excited radius (Figure 1.14, right). The excited ion packet induces a current when it orbits pass the detection plates. The induced current can be represented by the equation below:[33]

$$\frac{\partial Q}{\partial t} = -\frac{2q(\frac{\partial y}{\partial t})}{d} \qquad \text{[Eq. 17]}$$

where Q is the induced image current, q is the charge of the ion, y is the position of the ion on the y-axis, and d is the distance between the ion and the detection plate.

From equation 17, it is worth to note that the induced image current is independent to the magnetic field strength created by the superconducting magnet. Also, the position of the ion in the y-axis (y) is directly proportional the induced image current, meaning the detected ion signal increases with the cyclotron post-excitation radius. Moreover, the detected signal is also proportional to the charge of the ion (q), as a result FT-ICR MS (like all other MS platforms) is more sensitive to multiply charged ions.

As described by equation 13, the cyclotron frequency is inversely proportional to the *m/z* ratio of the ion, therefore ions with lower *m/z* ratio orbit at higher speed inside the ICR cell. The consequence of such fast-moving ions is the need to scan fast enough to detect the high frequency species. Ions having higher frequency than the sampling rate will not be detected because they will make multiple revolution between each sample event. As a result, the sampling rate has always needed to be set fast enough to keep up with the high frequency species. To determine the sampling rate of the detection event, it is important to determine the highest frequency desired, also known as the "Nyquist frequency" and to avoid aliasing effect and obtain an accurate detected signal of the ions, the sampling rate has to be at least twice as fast as the Nyquist frequency.

1.1.2.3.5 FT-ICR Cells

The FT-ICR cell is the mass analyser in all FT-ICR MS based experiments. However, there are different designs of FT-ICR cells and each of them is optimised to provide more consistent performance or to overcome challenges during FT-ICR MS experiments.

Cubic ICR cells were the first ever design for FT-ICR cells and have been used for many successful platforms over the years.[34] A cubic ICR cell consists of two

excitation plates, two detection plates, and two trapping plates, all identical in dimensions, able to fit effectively between the poles of permanent magnets (Figure 1.16a). However, cubic ICR cells possess limited ion capacity and were not optimised for cylindrical bore superconducting magnets which are more used in recent decades. Different cell geometries have been investigated over the years for better performance and cylindrical geometry is found to enhance the sensitivity of the experiment.

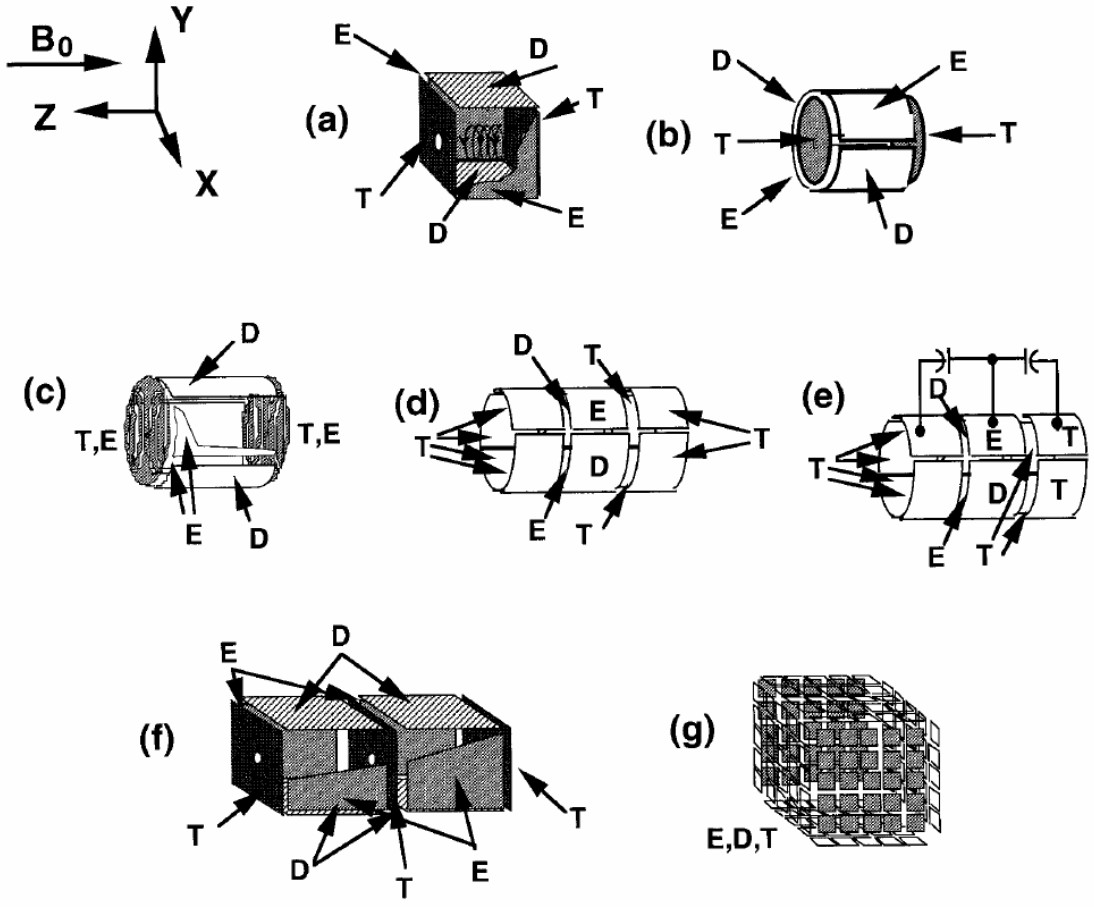


Figure 1.16 ICR cells configurations. E = excitation, D = detection, and T = trapping. a) Cubic ICR cell, b) Cylindrical ICR cell, c) Segmented end cap cylindrical ICR cell (Infinity Cell), d) Open cylindrical ICR cell, e) Open capacitively-coupled cylindrical ICR cell (Beu Cell), f) Dual cubic ICR cell, and g) "Matrix-shimmed" ICR cell. (Reproduced from Marshall *et al.* 1998 with permission).[29]

Cylindrical ICR cells can be classified into two main groups, closed and open. They both share a similarity of consisting rounded, quarter-circumference excitation/detection plates. Closed cylindrical ICR cells (Figure 1.16, b and c) use circular end cap trapping electrodes, while open cylindrical ICR cells (Figure 1.16, d and e) use cylindrical trapping plates for confining ions inside the ICR cells for measurement. The rounded/curved shape of the detection electrodes allow a more prolonged interaction with the ion-packets when they are excited, thus improving sensitivity (Figure 1.17). The end capped cylindrical ICR cells (Figure 1.16, b) suffer from more electric field permeation throughout the ICR cell and disruption of the ion orbit away from the centre of the ICR cell.[35]

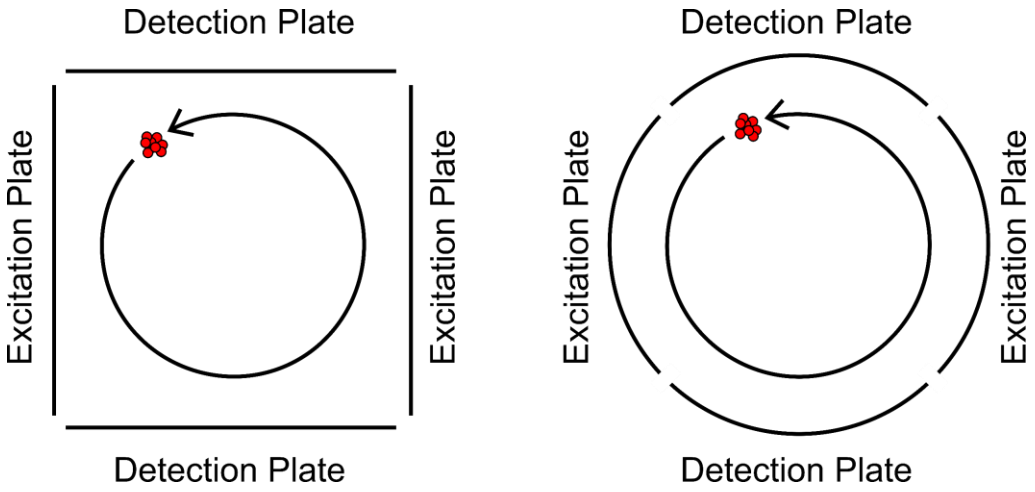


Figure 1.17 Ion packet interacts with the detection plate in the cubic ICR cell (left), less interaction when the ion packet is near the corners of the cubic cell while there is a more consistent interaction between the ion packet and the detection plate in the cylindrical ICR cell (right).

To overcome the problem caused by the end cap electrodes and a cylindrical cell of finite dimensions, Caravatti *et al.*[35] introduced the idea of a closed-cylindrical cell with special resistively-coupled trapping electrodes wire which was designed to simulate an infinitely long trapping electrode (Infinity cell, Figure 1.16, c), to eliminate axial excitation and ejection events. The capacitively-coupled open cylindrical cell (Beu cell, Figure 1.16, e) was introduced by Beu *et al.* in 1992.[36] Beu cell was designed to

capacitively-couple the excitation and detection electrodes during ion excitation to eliminate z-axis ejection of ions along the magnetic field.

1.1.2.3.6 Resolving Power of FT-ICR

The resolution of a peak detected in FT-ICR MS can also be calculated using the equation below:[37]

$$Resolution\ (R) = \frac{m}{\Delta m} = \frac{1.274\times10^{7} qBT_{aqn}}{m} \qquad \text{[Eq. 18]}$$

where m is the *m/z* ratio of the ion, q is the charge is the ion, B is the magnetic field strength, and T_{aqn} is the transient length.

From equation 18, the resolving power in a FT-ICR instrument is directly proportional to the magnetic field strength and the transient length. Ions inside the ICR cell ideally can orbit indefinitely, i.e. transient lasting indefinitely. However, practically ions experience ion-ion interactions and ion-background gas molecule collisions, disrupting the cyclotron frequency of the ions, and eventually ejected from the ICR cell, as discussed above. Therefore, it is important the pressure in the region where the FT-ICR cell sits remain as low as possible, ideally < 10^{-10} mbar. The magnetic field strength created by the superconducting magnet also contributes linearly to the resolving power too, *e.g.* from a 12 T FT-ICR mass spectrometer to a 21 T FT-ICR mass spectrometer, the resolving power should be able to increase by a factor of 1.75, for ions with the same *m/z*.

1.1.3 Tandem Mass Spectrometry

Mass spectrometry is a powerful tool which provides an insight into the mass information of the analyte, but it only provides an overall picture of which component is present in the sample mixture. Although elemental composition analysis, also known as fine structure, can be achieved by FT-ICR MS with the aid of its ultra-high resolution and high mass accuracy, it is often limited to small molecules. As the mass of the analyte increases, *e.g.* large biomolecules (proteins, DNA, etc.), fine structure analysis becomes extremely difficult as the mass increase can be a result of having different combinations of elements/components. Techniques such as electron impact ionisation

(EI, discussed in section 1.1.1.1) induce fragmentation during the ionisation process, which has its own advantages and disadvantages. With the fragments EI ionisation generates, structural information of the analyte can often be obtained, however, fragmentation during ionisation relies on pure/pre-separated samples to avoid overlap of fragment species from different analytes. To overcome this, selected fragmentation of species can be applied. Ions/precursors of interest can be selected/isolated by multipole mass selection (*e.g.* quadrupole mass selection/isolation), followed by applying different fragmentation techniques to yield important information from the fragment ions generated. This process is known as tandem mass spectrometry (MS/MS).

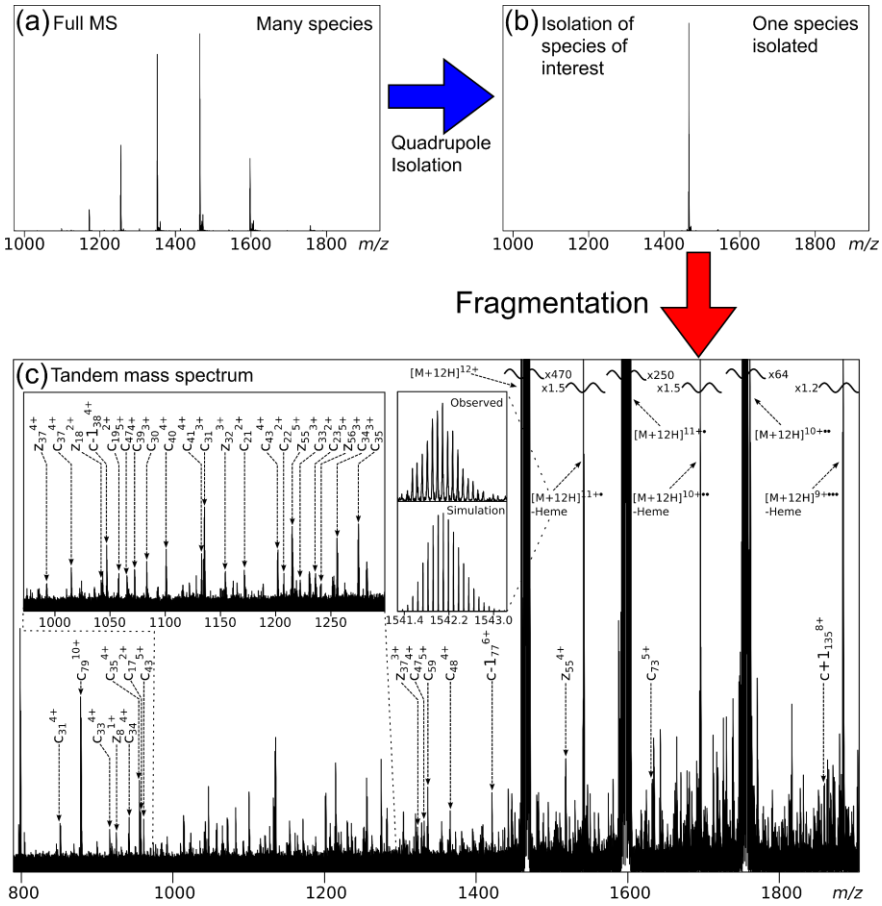


Figure 1.18 Standard acquisition in an MS/MS experiment. a) full mass spectrum of all species presents in a sample, b) isolation spectrum of a selected ion using a quadrupole mass filter, and c) tandem mass spectrum of the selected ion achieved by applying fragmentation. Sample shown in the MS/MS example is myoglobin from equine heart (a 16.9 kDa protein).

A wide range of fragmentation techniques have been developed over the years. These fragmentation techniques can be classified into three categories: collision-based, photon-based, and electron-based. Common fragmentation techniques include: collisionally activated dissociation/collisionally induced dissociation (CAD/CID),[38] infrared multiphoton dissociation (IRMPD),[39] ultraviolet photodissociation (UVPD),[40] electron capture dissociation (ECD),[41] electron induced dissociation (EID),[42] electron detachment dissociation (EDD),[43] electron transfer dissociation (ETD),[44] etc. In this section, three of the most common fragmentation methods, CAD, IRMPD, and ECD will be discussed. Peptides and proteins are the main focus in this thesis, a nomenclature system, also known as the "Roepstorff nomenclature", of fragments generating from a peptide/protein is summarised below (Figure 1.19):[45]

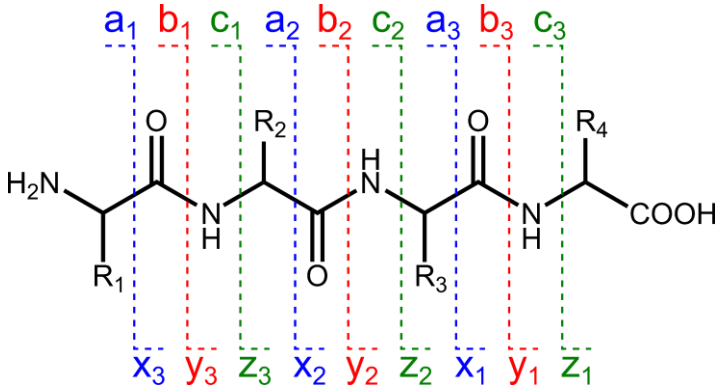


Figure 1.19 Roepstorff nomenclature of fragments generated from peptides/proteins.

1.1.3.1 Collisionally Activated Dissociation/Collisionally Induced Dissociation

Collisionally activated dissociation tandem mass spectrometry (CAD MS/MS) was first introduced by Jennings in 1968.[38] An inert collision gas (often Ar, Ne, He, or N_2) is involved in the fragmentation process. Analyte ions are usually accelerated into a cloud of collision gas by using an opposite polarity voltage. By forcing the ions of interest to collide with a cloud of neutral inert gas, it causes energy transfer and vibrationally heat up the ions of interest. Often a single collision event is not enough to heat up the ions of interest to fragment, but each ion will undergo many collision events, which allows the build-up of energy within ions. Eventually, the deposited energy will

exceed the bond strength of a bond within the analyte ion and it will dissociate and cause fragmentations. In peptides and proteins, the fragmentation often occurs at the kinetically weakest amide bond of the ion and generate b and y ions (see Figure 1.20).[46] CAD MS/MS does not only generate b and y ions in the fragmentation process, but also produces the loss of a wide range of small neutral molecules, *e.g.* H_2O, NH_3, CO_2 etc, which often increases the complexity of the mass spectrum and makes the data analysis more challenging.

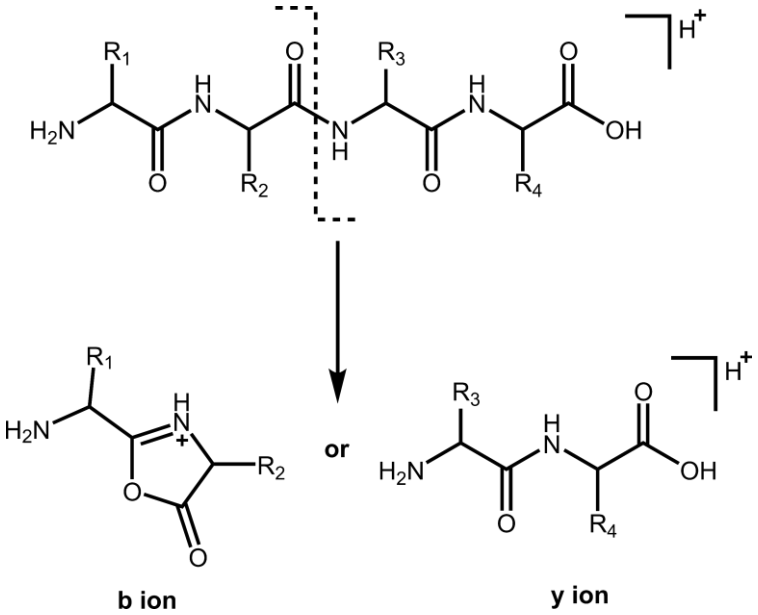


Figure 1.20 Diagram to show the breakage of the amide bond and generates b and y ions in CAD MS/MS.[46]

CAD MS/MS has been very effective at sequencing peptides, and has limited success at large peptides/small proteins, however, fragments are often only observed near the termini regions (N- and C- termini); rarely showing informative backbone coverage across the centre/the core regions of the protein.[47] As discussed above, CAD MS/MS breaks the weakest bond available in the analyte ion with slow heating generated from many collision events (ergodic process), therefore labile interactions/modifications, such as non-covalent interactions or phosphorylation will be dissociated during the CAD MS/MS fragmentation process and lose their location information in the analyte before backbone fragmentation, due to their inherently weak bond strength compared to the peptide backbone.[48]

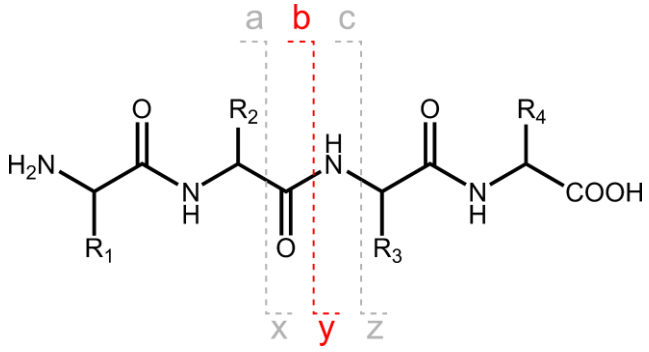


Figure 1.21 Representation of peptide backbone showing the bonds cleaved in CAD MS/MS and the types of fragment generated (highlighted).

1.1.3.2 Infrared Multiphoton Dissociation

Infrared multiphoton dissociation tandem MS (IRMPD MS/MS) is also a common technique that been used to study biomolecules, such as peptides and proteins.[49] Laser-based dissociation techniques have been shown to fragment the analyte ions effectively and achieving sequence information. Carbon dioxide (CO_2) laser is one the most common lasers used in IRMPD MS/MS because it is relatively cheap, continuous wave, and low maintenance. CO_2 laser emits photons at a wavelength of 10.6 μm which can readily be absorbed by many biomolecules and small molecules.[39] In an IRMPD MS/MS experiment, a 10.6 μm continuous-wave CO_2 laser is used to irradiate the trapped ions of interest (Figure 1.22), often in an ICR cell. Like CAD MS/MS, b and y ions are often generated upon IRMPD irradiation, often with losses of small neutral molecules, such as H_2O and CO_2 (Figure 1.23).

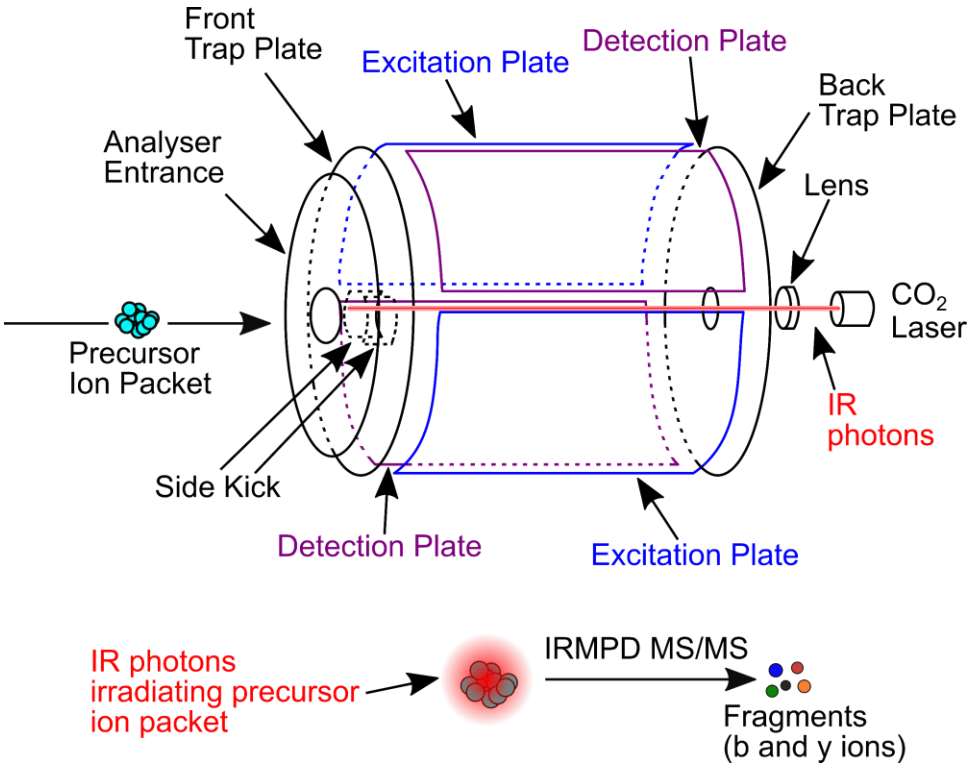


Figure 1.22 Schematic drawing showing the IRMPD MS/MS in the ICR cell. IR photons generated by the CO_2 laser irradiates the precursor ion packet which travels on the z-axis and getting fragmented in the b and y ions.

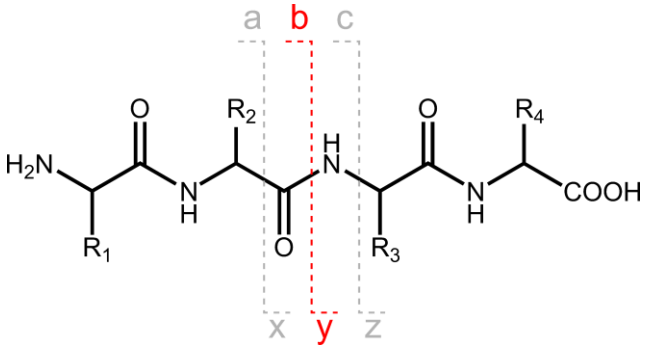


Figure 1.23 Representation of peptide backbone showing the bonds cleaved in IRMPD MS/MS and the types of fragment generated (highlighted).

Unlike CAD MS/MS, which fragments the analyte ions by colliding them with a cloud of stationary inert collision gas, IRMPD induces fragmentation by exciting the analyte ions into a higher energy state (excited state), overcoming the energy barrier of breaking the peptide backbone. Each photon generated by the CO_2 laser (~25 W)

provides only ~ 0.1 eV of energy, therefore, a dissociation event often requires an absorption of multiple photons. The level of fragmentation of CAD MS/MS can be optimised by the amount of collision energy applied, i.e. the amount of voltage applied to accelerate ions into the cloud of collision gas, while IRMPD MS/MS can often be optimised by the laser power and the pulse length of irradiation. The higher the power and the longer the pulse length of irradiation, the more the energy transfer to the analyte ions. Usually a careful balance between the laser power and pulse length of irradiation is required for an IRMPD MS/MS experiment, to avoid a complete dissociation of the analyte ions and a wide range of neutral losses.

1.1.3.3 Electron Capture Dissociation

Electron capture dissociation tandem MS (ECD MS/MS) was introduced by Zubarev *et al.* in 1998 to fragment multiply charged proteins.[41] ECD is usually implemented in FT-ICR MS while CAD and IRMPD are more commonly found implemented in other instruments as well. ECD MS/MS was believed to be a non-ergodic process where the precursor ions isolated (often by a quadrupole) is dissociated before the energy is distributed in the ions.[50] ECD is based on the emission of low energy electrons (~ 1 – 3 eV) and interacts with the trapped precursor ion packet. The electrons can be captured by positively charged analyte ions, creating a radical cation species (also known as the charge reduced species, CRS). The CRS are often energetically unstable and dissociates into multiple fragmentation pathways (discuss below). ECD MS/MS often results in the breakage of the N-C_α bond on the peptide/protein backbone and generates c and $z^\bullet$ ions as major products (Figure 1.24, a). Apart from c and $z^\bullet$ ions, a and y ions can also be generated as minor products in ECD MS/MS (Figure 1.24, b).

The electron capture event and the generation of the CRS can be represented by the equation below:[51]

$$[M + nH]^{n+} + e^{-}(\sim 1 - 3\ eV) \rightarrow [M + nH]^{(n-1)+\bullet} \qquad \text{[Eq. 19]}$$

where $[M+nH]^{n+}$ is the molecular ion with n protons.

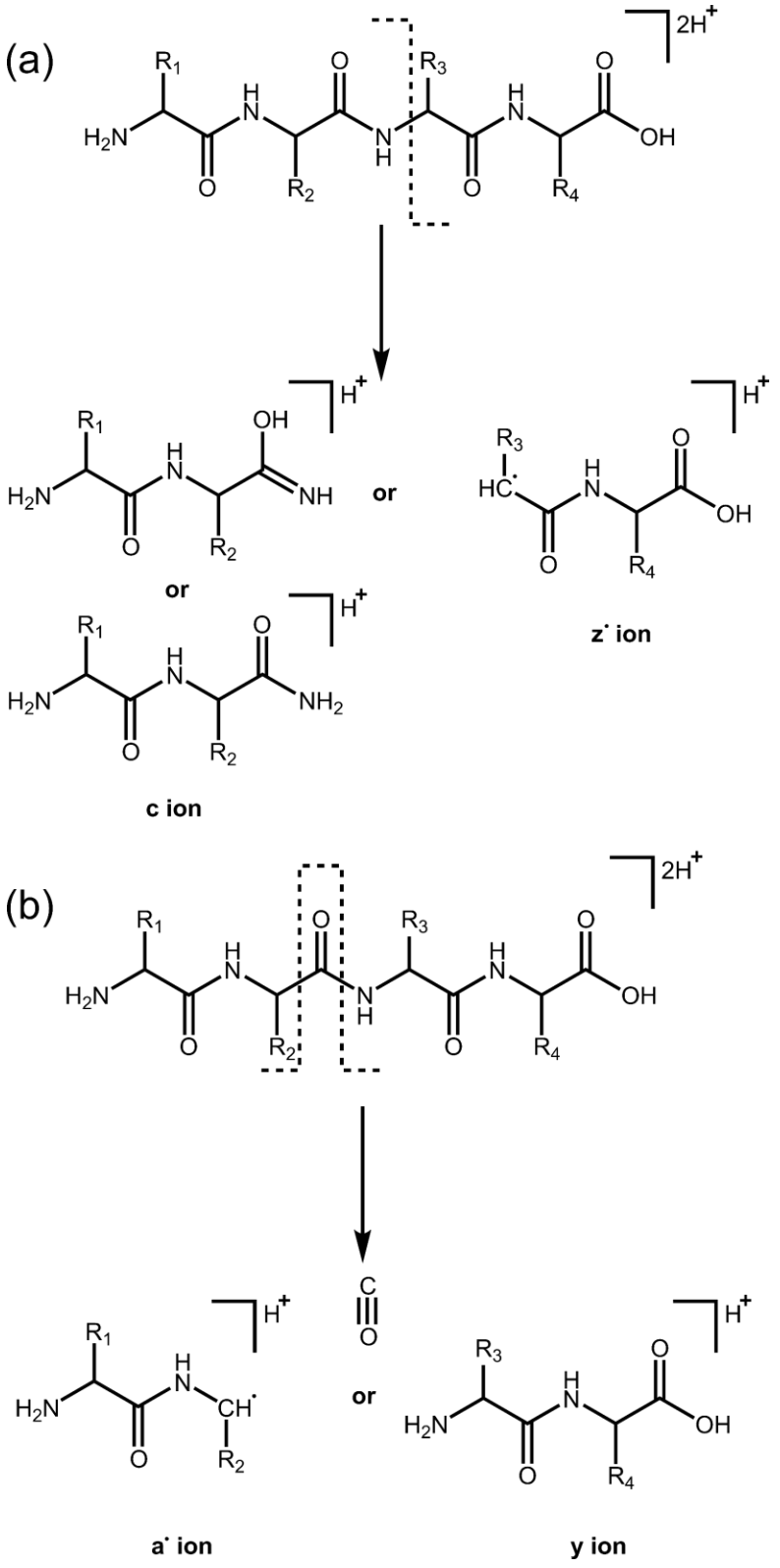


Figure 1.24 Reaction scheme to show a) the generation of c and z˙ ions as major products, b) the generation of a and y ions as minor products in ECD MS/MS.[46]

Often a hollow cathode is implemented behind the ICR cell to generate the low energy electrons for ECD MS/MS. In theory, ions are transmitted along the z-axis of the instrument, which unfortunately cannot interact with the cloud of electrons emitted by the hollow cathode as the cloud of electrons will be emitted as a "doughnut" shape, i.e. having no electron density in the middle, and most electron density form an outer circular ring (Figure 1.25). ICR cell optimisation (*e.g.* side kick voltages, trap plate voltages) is often needed to manipulate the selected precursor ion packet into the most electron-dense region to maximise the interaction between the precursor ions and the low energy electrons, thus yield more effective fragmentation from ECD MS/MS.

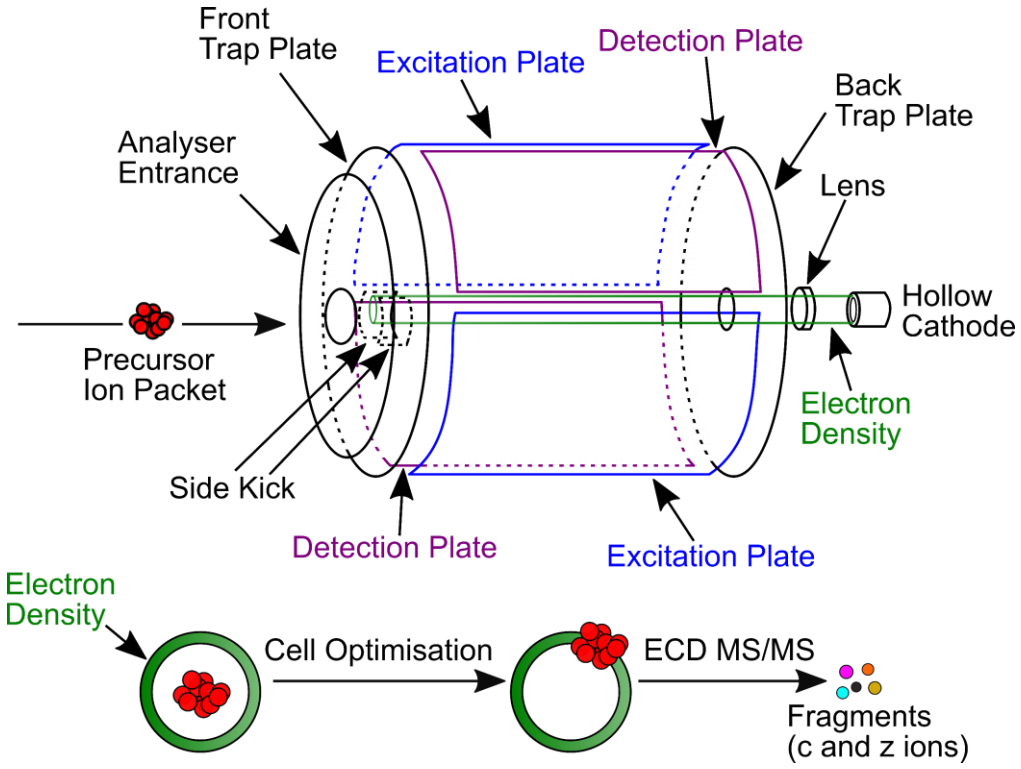


Figure 1.25 Schematic drawing showing the ECD MS/MS in the ICR cell (top). Low energy electrons are emitted from a hollow cathode. Ions are initially having no interaction with the cloud of electron, often cell parameters need to be optimised to drift the ion packet into the electron density to yield effective fragmentation, forming c and z ions.

The mechanism of ECD MS/MS in peptides/proteins has been under debate for many years, and there are a few more commonly accepted mechanisms in the field, below are the two major reaction mechanisms proposed for ECD MS/MS.

1.1.3.3.1 Cornell Mechanism

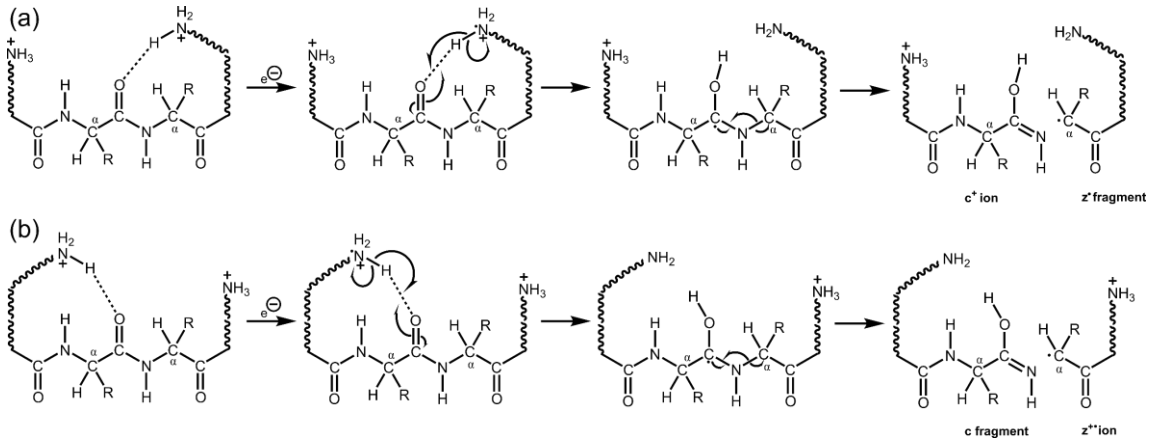


Figure 1.26 The Cornell mechanism of ECD MS/MS. a) electron capture at the C-terminal donor amine group, causing the breakage of N-C_α and generates c^+ and $z^\bullet$ ions, b) electron capture at the N-terminal donor amine group, causing the breakage of N-C_α and generate c fragment and $z^{+\bullet}$ ions.[51]

The Cornell Mechanism was first introduced by McLafferty *et al.* in 1998.[41] Electron capture event occurs at a positively charged site (protonated N-terminus or on the side chains of basic amino acids (lysine (Lys), arginine (Arg), histidine (His))), forming a hypervalent N-species, inducing a hydrogen transfer to the amide oxygen, resulting in the formation of a carbon-centred aminoketyl radical. The breakage of the N-C_α bond is then induced by being physically next to an unstable radical species and forming c- and z- ions (Figure 1.26).[51]

1.1.3.3.2 Utah-Washington Mechanism

The Utah-Washington mechanism is a combination of two previously proposed mechanisms: the Utah mechanism by Simons *et al.* in 2003,[52] and the Washington mechanism by Turecek *et al.* in 2005.[53] The two mechanisms focused on different steps in the dissociation event: Utah focused on the electron capture steps while Washington focused on the subsequent dissociation. However, the two mechanisms proposed the

formation of same products (c- and z$^{\bullet}$- fragments from the breakage of the N-C$_\alpha$ bond in peptides/proteins). Therefore, the two mechanisms are combined and known as the Utah-Washington mechanism and is the most accepted mechanism for ECD MS/MS (Figure 1.27).

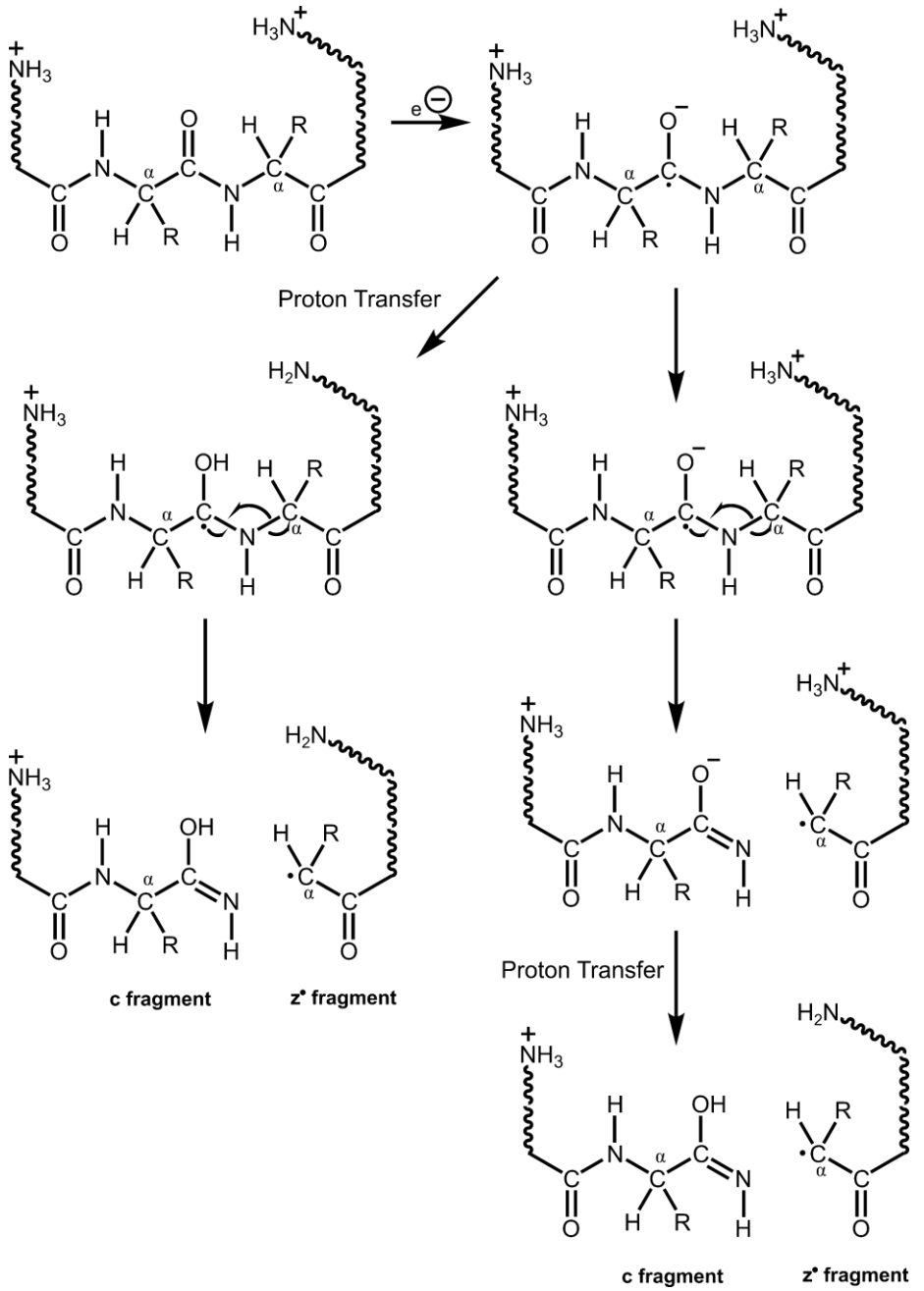


Figure 1.27 The Utah-Washington mechanism of ECD MS/MS, showing the two possible pathways to dissociate the N-C$_\alpha$ bond after the electron capture event.

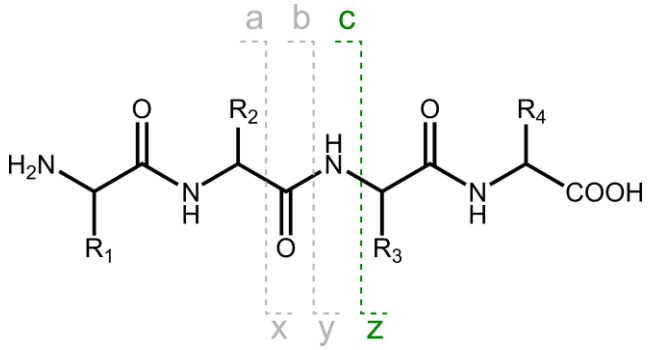


Figure 1.28 Representation of peptide backbone showing the bonds cleaved in ECD MS/MS and the types of fragment generated (highlighted).

1.1.3.3.3 ECD of Disulfide Bonds

In CAD and IRMPD MS/MS, fragmentation effectiveness is often hindered by the number of disulfide bonds present in the sample, the number of possible cleavages along the peptide/protein backbone is less with CAD and IRMPD MS/MS as often the disulfide bonded region is protecting the backbone to be cleaved (Figure 1.29). While ECD MS/MS is shown to preferentially cleave disulfide bonds in a sample and allow more possible cleavages.[54]

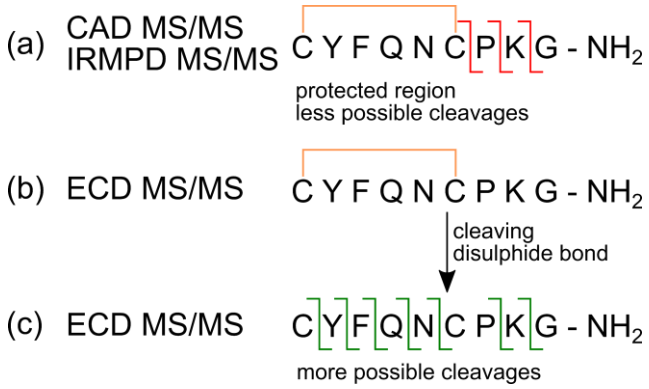


Figure 1.29 Example of the effectiveness of different fragmentation techniques on a disulfide bonded sample. a) limited possible cleavages (shown in red) with CAD or IRMPD MS/MS due to the protected region by the disulfide bond in the sequence (shown in orange), b) disulfide bond is getting cleaved by ECD, opening the protected region, c) more possible cleavages (shown in green) without the protection by the disulfide bond. Example was shown using Lys^{8} – vasopressin.

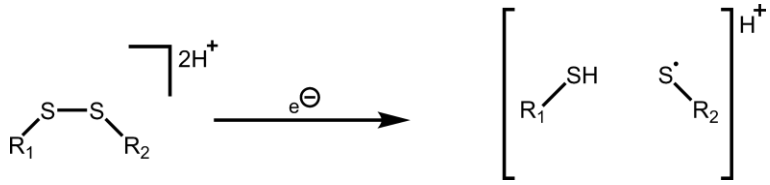


Figure 1.30 Reaction scheme showing the capture of electron to the a disulfide bond and results in S-S bond cleavage.[46]

1.1.3.3.4 Fragmentation Efficiency of ECD MS/MS

Efficiency of a fragmentation technique can be represented by the equation below:

$$Fragmentation\ Efficiency = \frac{\sum A_{fragments}}{\sum A_{precursor}} \qquad \text{[Eq. 20]}$$

where $\Sigma A_{fragments}$ is the sum of the peak areas of fragment ions after MS/MS and $\Sigma A_{precursor}$ is the sum of peak areas of precursor ions before MS/MS.

The fragmentation efficiency of ECD MS/MS is ~ 20% for a doubly charged peptide ion, however, a lot higher fragmentation efficiency, as high as 80% efficiency can be achieved with some highly charged protein ions.[55] To enhance the efficiency of ECD MS/MS in some highly structured proteins, often IRMPD is applied to aid the fragmentation process. The IRMPD applied is often tuned to induce as minimal fragmentation effect as possible, it is mainly to warm up the precursor ions, activating the structure to become more accessible by electrons during ECD MS/MS. This process is known as infrared-activated electron capture dissociation tandem mass spectrometry (IR-ECD MS/MS).[56,57]

1.1.3.3.5 Applications of ECD MS/MS

ECD MS/MS has shown very promising fragmentation in top-down analysis, produces extensive cleavages along the protein backbone.[41] Therefore, ECD MS/MS is an excellent tool to perform *de novo* sequencing, i.e. to determine the peptide/protein sequence by using fragment ion information from a tandem mass spectrum. *De novo* sequencing is extremely useful if an unknown sample is involved, by determining the identity of the unknown component in the sample, often expanding the knowledge to the

study significantly. Multiple fragmentation techniques can be applied to a sample to maximise the backbone cleavages, by combining them together, they contribute hugely into the *de novo* sequencing of biomolecules.

CAD and IRMPD MS/MS cleave the peptide/protein backbone by slowly heating up the precursor ions by energy transfer *via* multiple collision events and exciting the ions to their excited states using IR photons, respectively, which are ergodic processes. Therefore, these fragmentation techniques tend to lose the important labile post-translational modifications (PTMs), such as phosphorylation[58] and glycosylation[59], and non-covalent interactions[60] before breaking the peptide/protein backbone. This restricts the possibility of studying the modification/binding locations on biomolecules. While ECD MS/MS is a radical-based process, which is capable to preserve these labile modifications and non-covalent interactions, allowing the precise modified locations or binding sites to be derived.

1.1.3.3.6 Other Electron-Based Dissociation Techniques

Apart from ECD, there are other electron-based dissociation techniques (ExD) that are employed to work on different sample natures. Electron transfer dissociation (ETD) was introduced by Syka *et al.* in 2004.[44] Unlike ECD using a cathode to generate electrons, ETD relies on a carrier molecule/ion to transfer the electron to the analyte species. Negative anion radicals, generated by a negative chemical ionisation (nCI) source (Figure 1.31), and the precursor ions are transferred to the ion trap MS where ion-ion interaction takes place. During the ion-ion interaction process, electrons could transfer to the analyte species and create CRS, and eventually produce c- and z-fragment ions like ECD.

Instead of using low energy electrons to induce fragmentation like in ECD, higher energy electrons can also be used. Electrons with energy (~ 6-10 eV) can interact with the precursor ions, not only generating the c- and z- ions, but also special side chain losses from amino acid residues. Some of these side chain losses are diagnostic, with observation of these side chain losses, revealing the presence of specific residues. The use of electrons within this energy range is referred as hot-ECD (hECD).[51]

ECD has proved to be a very effective fragmentation technique, however its use is limited to multiply charged precursor ions. Small molecules are often observed as singly charged ions, and they cannot be fragmented with ECD due to the charge limitation. Therefore, electron induced dissociation (EID) was developed by Budnik *et.al.* in 2000.[61] Electrons within the energy range of ~ 9-25 eV can be used to perform EID, generating fragment ions similar to those from EI.

ECD is an exclusively positive-mode technique, meaning it cannot be used on negatively charged ions. Therefore, electron detachment dissociation (EDD) was introduced by Budnik and Zubarev *et al.* in 2001.[62] Electrons within energy range of 15-20 eV could induce electron detachment on the negative target analyte, therefore, EDD is a common fragmentation technique to study DNA/ribonucleic acid (RNA) systems.[43]

1.1.4 A Hybrid FT-ICR Mass Spectrometer

All MS data presented in this thesis are acquired using a Bruker 12 T SolariX FT-ICR MS (Figure 1.31). Ions can be generated using ESI, nESI, or MALDI in a SolariX FT-ICR MS. Ions produced by ESI and nESI are orthogonally induced *via* a glass capillary into the vacuum region and accelerated towards to ion funnels by applying an equal polarity voltage to the incoming ions onto the deflector plate, this allows neutral molecules to be taken away from the vacuum while ions are transmitted further down into the FT-ICR MS. Ions produced using MALDI are generated under vacuum under a MALDI extraction lens, which accelerates the ions through the deflector plate and getting repelled towards the ion funnels.

Ions can be focused using ion funnels set at the entrance of the multipoles. Focused ions are then transmitted into the Qh (quadrupole-hexapole) region of the MS. The quadrupole can be used for transmission purpose only or for mass filtering if MS/MS experiment is required. Ions of interest can be selected in the quadrupole (discussed in section 1.1.2.1) and transmitted further down into the hexapole collision cell for ion accumulation. CAD or ETD can also be performed in the hexapole collision cell if needed.

Ions are then transferred into the ICR cell through a hexapole transfer optic. Once the ions entered the ICR cell, which sits in a homogenous magnetic field region within the superconducting magnet, by applying a voltage to the front trap plate (equal polarity to the ions) and a similar magnitude to the back trap plate (usually 0.35-1 V for infinity cell), ions can be confined in the ICR cell.

Once ions are confined in an ICR cell, IRMPD and ExD can be performed if necessary. The hollow cathode and CO_2 laser are located at the opposite side of the ion entrance in the ICR cell with lenses positioned between the cathode and the ICR cell. All ions can be excited and detected inside the ICR cell as discussed above.

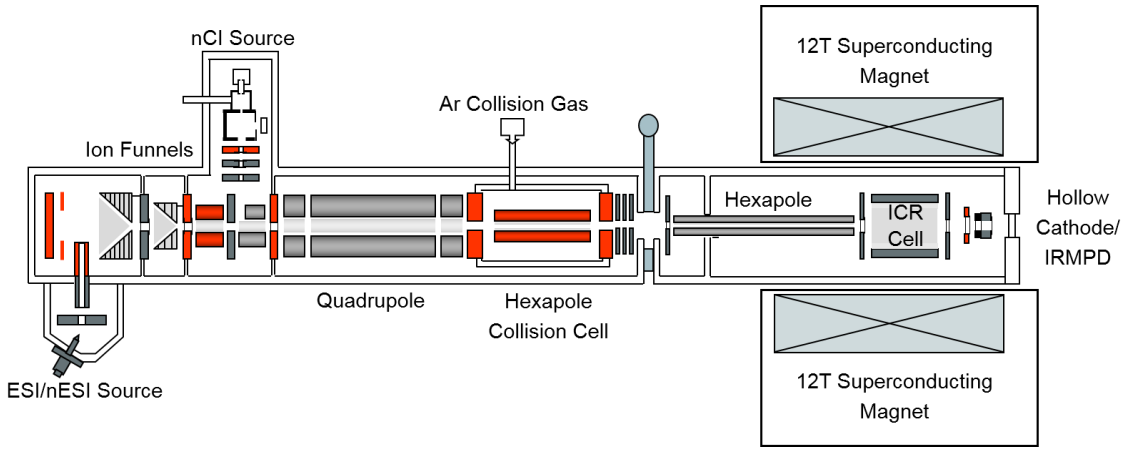


Figure 1.31 Schematic representation of the Bruker 12T SolariX mass spectrometer. Courtesy of Bruker Daltonics, Bremen, Germany.

1.2 Introduction to Anti-Cancer Metallodrugs

1.2.1 Cancer

Uncontrolled/excessive cell proliferation (growth and replication) is detrimental to individual cell function and is a common fundamental symptom of cancer. Growth and replication are often carefully controlled and limited by cellular proteins and biological mechanisms to sustain a healthy condition of a living organism. Inside the nucleus of cells, there are DNA coding information for cell growth and replication. Cancer often arises because of damages/mutations/coding problems within the DNA of the cell itself. Apart from chemical-induced damages (*e.g.* host exposure to carcinogens), radiation-induced damage (*e.g.* prolonged exposure to UV light),[63] or problem arising during DNA replication/transcription.[64] Damage to the host cell DNA can lead to an accelerated proliferation rate, demanding more resources from the host organism and creating asymmetric cell growth in that area/region of the host, often forming a large collection of cancerous cells, also known as a tumour (Figure 1.32) Cancerous cells can often migrate from the creation area/primary site to other areas of the host, creating secondary tumours, a process known as metastasis (Figure 1.33). The spread of cancerous cells can put a huge amount of pressure/strain on biological processes, both from resources competition between the healthy cells and cancerous cells due to the accelerated proliferation and physical stresses from the cancerous cells/ the tumour to the cells nearby, disrupting normal cell functions. Unfortunately, the probability of developing cancer is unavoidable as there is always a possibility of DNA damage whenever a cell undergoes division, creating damaged DNA.[65] The development of cancer and its biological causes have been the subject of detailed studies over many years.[66]

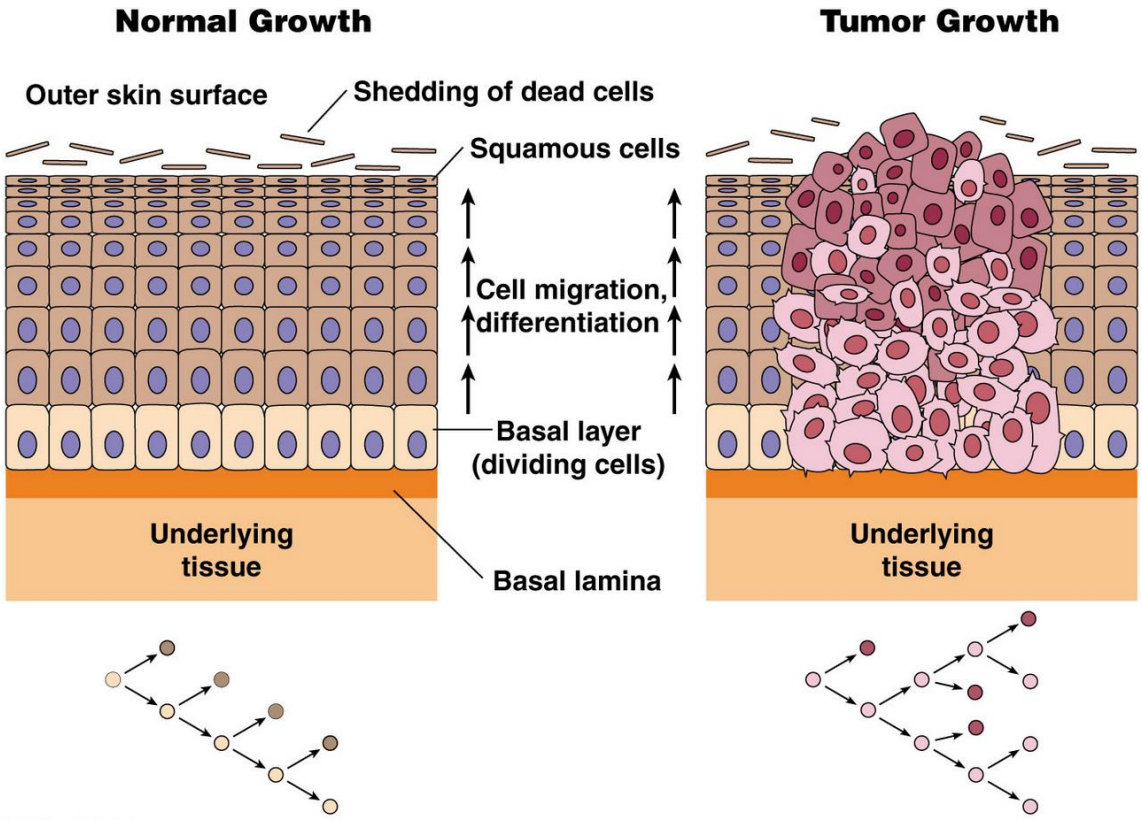


Figure 1.32 Normal and uncontrolled cell growth after DNA damage. An unbalanced growth of cancerous cells (tumour) is competing for resources and putting stress on normal cells nearby. Reproduced from http://www.mun.ca/biology/desmid/brian/BIOL2060/BIOL2060-24/24_01.jpg, accessed 20/05/2019.

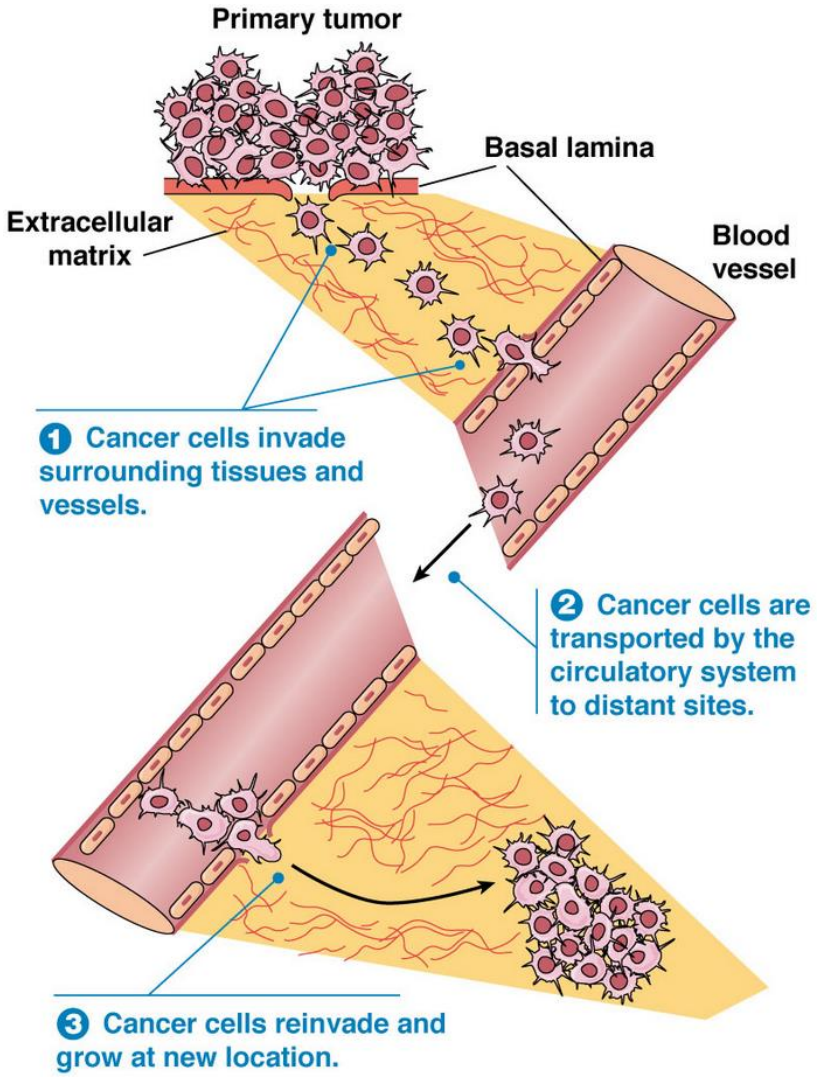


Figure 1.33 Primary tumour migrates from the creation point to distant sites by the circulatory system, forming a secondary tumour. Reproduced from http://www.mun.ca/biology/desmid/brian/BIOL2060/BIOL2060-24/24_04.jpg, accessed 20/05/2019.

1.2.2 Cancer Treatment

Different forms of treatment have been developed and chemotherapy is one of the most commonly used approaches to attack cancer cells. An ideal cancer treatment would have high selectivity towards cancerous cells and extremely low or no selectivity towards normal healthy cells.[67] Potencies of anti-cancer drugs can be measured by their half-maximal inhibitory concentration (IC_{50}) values. The IC_{50} value is a measure of how

much drug is needed to inhibit a biological process by half, thus the smaller the IC_{50} value of a drug, the more potent the drug is for that process.[68] These treatments are often designed to target DNA as cancer is intrinsically DNA-related. Different DNA-targeting reagents have been designed over years, and are usually small molecule-based,[64] macromolecule-based,[69] and transition metal-based.[70,71]

Ideally, these metal-based anti-cancer complexes target cancerous DNA within the cancer cells and induce a chemical change, often *via* covalent bonding, crosslinking, or deformation of structure.[72,73] The "modified" DNA will no longer able to be interpreted by transcription proteins; preventing the cancerous cell from further replicating, and eventually leading to cell death. However, because of the highly similar cell content and cell function between the normal host cell and cancerous cells, a perfect anti-cancer drug is extremely difficult to design, i.e. high selectivity to cancerous cells and non-potent to normal cells. Metal-based anti-cancer drugs are the most commonly used chemotherapy. Over the past few decades, platinum based anti-cancer drugs made a huge impact on cancer therapy. The first Food and Drug Administration (FDA) approved platinum-based drug was cisplatin (*cis*-diaminodichlorido-platinum, CDDP).[74] Cisplatin is shown to attack cancerous DNA by physically binding to the exposed guanine nucleobases, causing DNA deformation, disrupting its tertiary structure, preventing transcription and proliferation and leading to cell death (Figure 1.34).

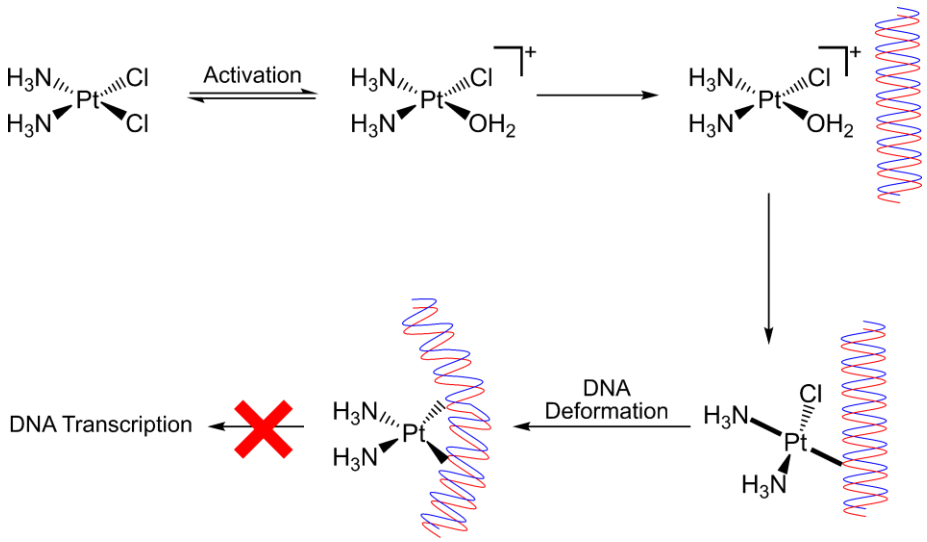


Figure 1.34 Mechanism of action of cisplatin on DNA. Cisplatin is activated inside the host's body and bind to the DNA, which causing DNA deformation and fail to be transcribed.

Cisplatin has shown its effectiveness against various different cancers, especially testicular, cervical, and bladder cancers. Cisplatin has high potency, having IC_{50} of 1.2 μM (± 0.2) against A2780 ovarian cancer cells.[75] However extensive use of cisplatin on cancer patients has results in an array of side effects, including liver failure, kidney damage, dramatic weight loss etc.[76] With the widespread cisplatin treatment, cisplatin resistant strains of cancer have also been observed unfortunately. Due to the self - protection system of the host, any "toxic" treatments, including the use of cisplatin to attack cancer cells, the host will attempt to adapt and remove the active complexes from the host body, lowering the effectiveness of the drug. To overcome this, other platinum-based anti-cancer drugs, including oxaliplatin, carboplatin etc. have been developed.[74]

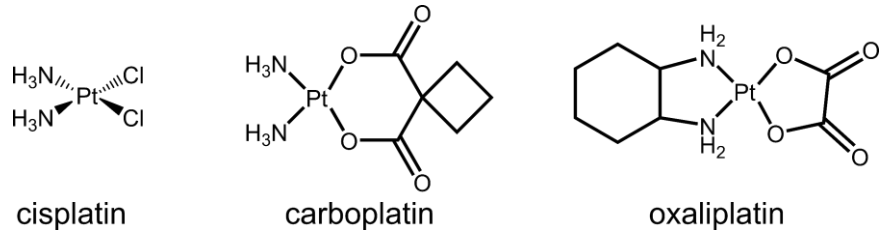


Figure 1.35 Examples of the traditional platinum-based anti-cancer drugs.

Studying the mechanism of action (MoA) of a drug is important because it allows researchers to understand better how the drug works and provides insight into future drug development. For instance, cisplatin has shown that one of more of the chlorido ligands in the structure can be displaced while inside the host body/cell, producing a more active aqua-Pt(II) complex (Cl^- displaced by H_2O), and this aqua/di-aqua species can bind to cancerous DNA to cause structural deformation, preventing further cell growth and replication of cancerous cells, inducing cell death (Figure 1.34).[77,78] However, even cisplatin has proven to be very effective against certain cancers, only a very small percentage of the cisplatin actually makes its way to the DNA target in the cancerous cells.[79,80] There are far more species/molecules in the host body which could bind to cisplatin before the drug makes its way to the DNA target despite the guanine nitrogen N7 presenting a very reactive ligands for cisplatin. Studies have shown 90-98% of cisplatin actually becomes protein-bound with just one day of injection.[79,80] It is believed that the undesired side effects experienced by patients are due to the off-target interactions. Therefore, the design of novel metallodrugs which have fewer side effects and can overcome resistance to current platinum drugs has attracted significant attention.

1.2.3 Novel Metallodrugs

In-depth studies of novel metallodrugs have been a popular research topic, including rational drug design and further development of the existing drugs. Below are a series of new metallodrugs, categorised according to their structure and functions, each with the own target/mechanism of action, designed to cause either cell death or prevention of them from proliferation.

1.2.3.1 Piano-Stool Complexes

In older metallodrug designs, multiple binding sites are often found on complexes by losing/displacing the different initial ligands in the complex structure, making the study of mechanism of action difficult. Therefore, slight alternations in drug designs have been proposed to restrict the loss of multiple ligands by introducing large chelating ligands (using more than one pair of electrons to bind to the metal ion centre). These large chelating ligands can be π-bonded aromatic rings as in arenes and

cyclopentadienyls (Figure 1.36), which bind to the metal ion centre in an η^6 or η^5 fashion, i.e. using 6 electrons and 5 electrons, respectively.

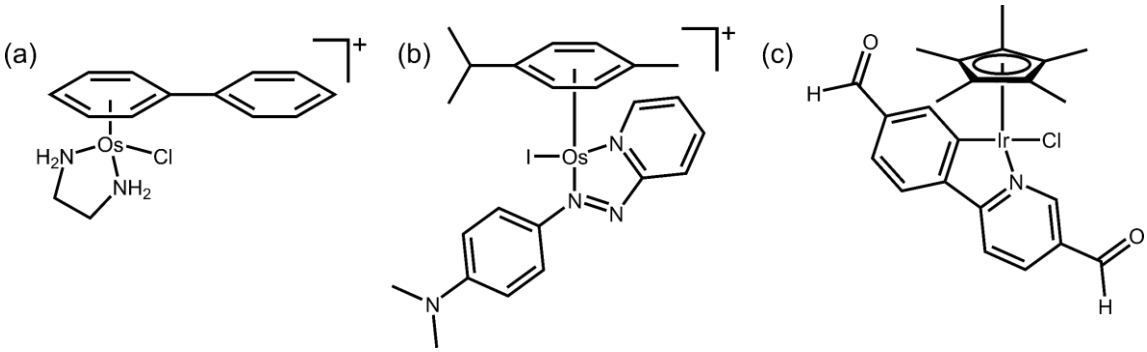


Figure 1.36 Examples of piano-stool complexes. a) [(η^6-bip)Os(en)Cl]PF_6,[43] b) [(η^6-*p*-cym)Os(NMe_2-azpy)I]PF_6,[81] and c) [(η^5-Cp*)Ir(2-(CHO-phenyl)-(CHO-pyridine))Cl].[60,82]

With using an η^5 or η^6-arene and a bidentate ligand, it was shown to greatly enhance activity at the monodentate site, making these piano-stool complexes excellent candidates to be active anti-cancer agents.[83–85] A recent study by Sadler and co-workers has reported an osmium-based piano-stool complex which was 49-times more potent than cisplatin towards A2780 ovarian cancer cells.[81] Though piano-stool complexes can be more active than cisplatin towards target cancerous cells, they are still often transformed *in vivo* from the initial complex to a more active species. This process often is involved in intelligent drug design: by introducing a relatively inactive species (pro-drug) to the host and only transform/activated when the pro-drug has travelled to the target area, this can be induced by a change in micro-environment (*e.g.* pH, temperature etc.).[70,73]

Selectivity is a key factor of drug design and development, to reduce side effects to patients during treatment and more importantly to enhance the drug selectivity towards the correct target and this can be achieved by employing functionalised ligands. The functionalised ligands are designed to bind to specific targets, leading to accumulation of drugs in certain areas, not only enhancing the selectivity of the drug but also make the drug more potent.[86]

1.2.3.2 Photoactivatable Complexes and Photosensitisers

As discussed above, high selectivity is a key factor for anti-cancer drug development. Therefore, a variety of metallodrugs have been designed to be activated by external stimulus. Apart from changes in micro-environment (pH and temperature), light can also be used as a trigger to transform the pro-drug to its active form.[26,87] Ideally, before light irradiation/activation, the pro-drug should possess little to no cytotoxicity (the quality of being toxic to cells), which means the pro-drug will not bind/interact with any molecule or cells, and upon irradiation, the pro-drug can be activated directly to have high cytotoxicity near/in the cancer cells/tumour. Obviously, if a pro-drug can be activated by light at all wavelengths, it will easily be activated before the drug reaching the desired target, losing the aim of being an effective photoactivatable drug. Therefore, it is important for a pro-drug only be activated at a specific wavelength, so that it can transform into its active form, creating an effective and controlled chemotherapy treatment, also known as photoactivated chemotherapy (PACT).[87] These photoactivatable complexes are usually activated by higher-energy wavelengths (*e.g.* blue visible light, ~ 450-490 nm), resulting in metal-ligand bond cleavages, leaving a metal complex of different geometry (*e.g.* octahedral to square planar),[26] which is often more active than the pro-drug, allowing it to attack the cancerous cells more effectively.

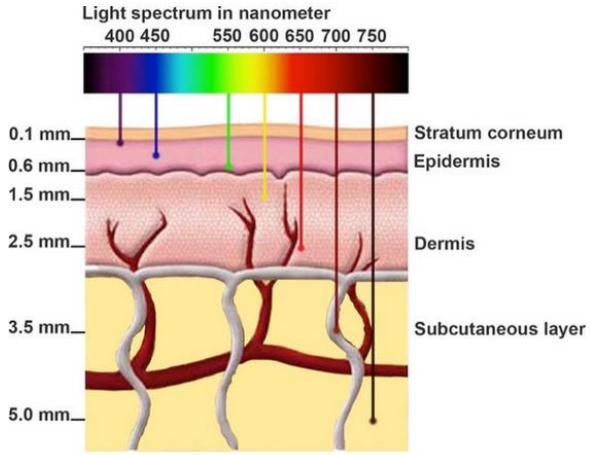


Figure 1.37 Light penetration measurements in millimetres through human tissue. (Reproduced from Wang *et al.* with permission).[88]

One of the leading PACT agents, *trans, trans, trans*-[Pt$(N_3)_2(OH)_2$(pyridine$)_2$], also known as FM190, has shown to be an effective PACT agent by having very low cytotoxicity of > 212.3 μM in the dark towards OE19 (oesophageal cancer) and upon irradiation with 420 nm, IC_{50} of FM190 is 8.4 μM.[89] Wootton *et al.* have recently shown FM190 is able to perform a dual attack on peptide amino acids, either by oxidation (induced by hydroxyl radicals), or platination (binding to the platinum centre), or both.[26]

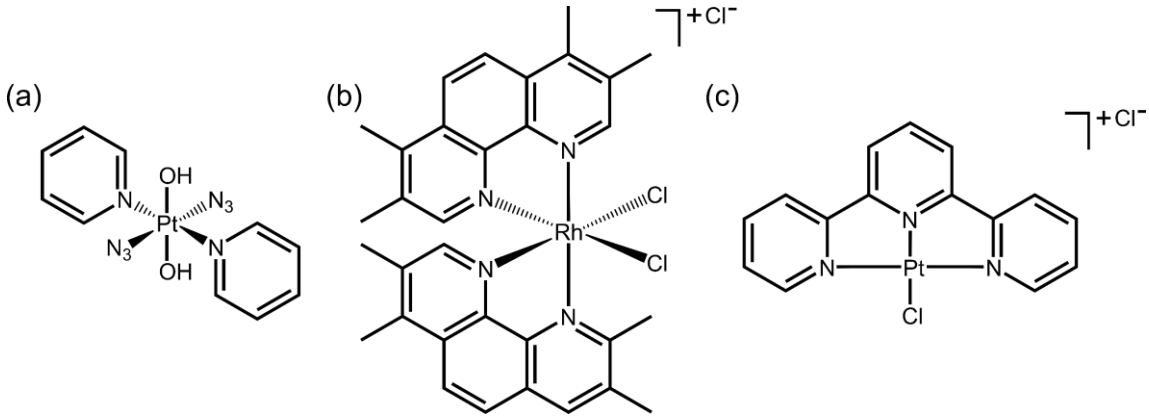


Figure 1.38 Examples of PACT agents. a) *trans, trans, trans*-[Pt$(N_3)_2(OH)_2$(pyridine$)_2$],[26] b) a methylated cis-[Rh(phen$)_2Cl_2$]Cl,[90] and c) [Pt(terpy)Cl]Cl.[91]

Photosensitisers (photodynamic therapy agents, PDT agents) for cancer therapy usually function by absorbing a convenient wavelength of light, often a good transmission/penetration through skin/tissue, but at the same time cannot be harmful to the host, *e.g.* visible light. Energy absorbed is then transferred to nearby cellular components, *e.g.* mitochondria, where oxygen molecules can be excited to become singlet oxygen ($^1O_2$). Singlet oxygen a type of reactive oxygen species (ROS) and is a much more reactive form of molecular oxygen. Having a high concentration of singlet oxygen within the target cell, can induce oxidative stress leading to cell death/apoptosis.[92] Zhang *et al.* have shown that the photosensitiser, [Ir(phenylquinoline$)_2$(diketonate$)_2$], also known as PZ6, being extremely potent towards A549 lung cancer cell line upon irradiation.[93] PZ6 is non-toxic in the dark, having IC_{50} value > 100 μM, while upon blue light (465 nm) irradiation, PZ6 is extremely potent

(IC_{50} = 0.23 μM). PZ6 has shown the ability to generate singlet oxygen upon irradiation, supporting by the observation of 3-hydroxykynurenine when a model peptide (Lys^3-Bombesin) was treated with PZ6.[94] Zhang *et al.* have also shown specific oxidative attack on specific cellular proteins when A549 are treated with PZ6.[93]

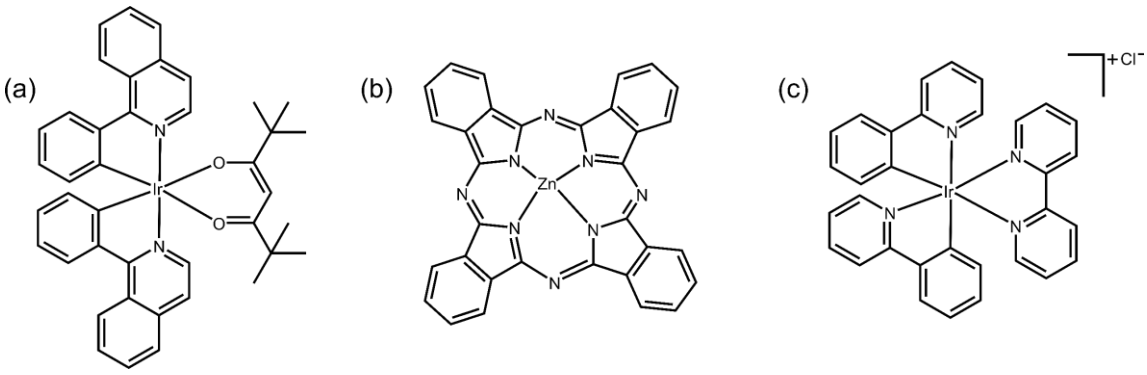


Figure 1.39 Examples of PDT agents. a) [Ir(phenylquinoline)$_2$(diketonate)$_2$],[93] b) Zinc phthalocyanine,[95] and c) [(2-(phenyl)(pyridine))$_2$Ir(2-(pyridine)$_2$)]Cl.[96]

1.3 Introduction to Proteomics

The terms "proteome" and "proteomics" were first introduced by Marc Wilkins in 1995.[97] Proteome is a combination of the words "protein" and "genome", meaning the total protein complement of a genome. Proteomics is the study of the proteome (a large scale of proteins generated by living organisms).

MS is a common analytical technique used in proteomics. By measuring the masses of proteins and peptides, information such as protein identities, modifications, including PTMs, or even structures of proteins can be revealed. Proteomics is important because it provides an insight into the mechanisms and inner working of cells,[98] which helps us to understand how the anti-cancer drugs are affecting cells, *e.g.* change in protein quantity or even a change in protein content. Proteomics research is hugely involved in various areas: biomarker discovery for diseases,[99] protein quantification,[100,101] PTMs study on proteome scale,[93,102] monitoring the interactions of potential drugs with proteins in living organisms.[93,103] MS-based proteomics can generally be classified into three different types: top-down proteomics, bottom-up proteomics, and middle-down proteomics, which will be discussed more below.

1.3.1 Top-Down Proteomics

Top-down proteomics is a study of intact proteins, it preserves the most information of the initial biological system, including intact protein sequence information and chemical modifications (*e.g.* PTMs) as less chemical alternation is needed to be induced to analyse the samples, which could disrupt labile modifications and interactions in-between or within proteins. Therefore, top-down proteomics analysis is normally used for structural analysis, PTM, and *de novo* sequencing studies of proteins.[57,101,104,105] Because of the huge number of proteins present in a proteomics sample and the possible fragments generated by tandem mass spectrometry (MS/MS), the performance requirement of the MS instrument is also the highest. Ultra-high resolving power is generally needed because of the overlapping distributions of peaks throughout the whole mass range which reduces the type of MS which can be used. FT-ICR MS and orbitrap MS are the two most common MS system required to fulfil the extremely high demand in resolving power. However, such instruments are very expensive and therefore top-down proteomics is not as commonly seen as the other two types of proteomics studies; bottom-up proteomics and middle-down proteomics.

1.3.2 Bottom-Up Proteomics

Bottom-up proteomics is the study of digested proteins. Proteases such as trypsin, pepsin, and chymotrypsin are used to break large proteins into manageable size peptides. Trypsin cleaves proteins at the carboxyl side of arginine and lysine while pepsin and chymotrypsin prefers cleaving at amino acids containing an aromatic ring, such as phenylalanine, tryptophan, and tyrosine.[106] Proteins that underwent digestion will be chemically cut into smaller pieces of peptides. However, before protein digestion takes place, chemicals such as reducing and alkylating reagents are commonly mixed with the protein samples to break disulfide bonds and help achieve a better result of digestion.[107] Dithiothreitol (DTT), and tris(2-carboxyethyl)phosphine (TCEP) are commonly used to reduce the disulfide bonds, opening the protein structure. Disulfide bonds will be reformed from the free thiols on cysteine residues though, therefore, alkylating reagents, such as iodoacetamide (IAA), N-ethylmaleimide (NEM) are used to react with the free thiols and block them off, preventing disulfide bonds from reforming.

After the reduction and alkylation steps, there should not be any cross-linking by disulfide bonds left in proteins, and the proteases can perform a more effective digestion. Digestion is often carried out under 37 ^{o}C, to simulate the body temperature, therefore, incubation is needed. Digestion time is often around 12-16 hours, with the long incubation time, PTMs such as oxidation and deamidation are often artificially induced into the sample and fragile PTMs could be lost during the process too. The bottom-up approach is relatively easier than top-down analysis because peptides produced from the digestion steps are easier to be detected and analysed compared to large proteins.

1.3.3 Middle-Down Proteomics

Middle-down proteomics is an approach which sits between top-down and bottom-up. Intact proteins are still being digested, but the product peptides are larger. Proteins are usually enzymatically digested with a shorter incubation time compared to the bottom-up approach (2-4 hours), this helps to reduce the loss of fragile modifications (*e.g.* PTMs) and decrease the amount of modification artificially induced into the samples. Middle-down digestion can also be achieved by employing different proteases. Endoproteinase Lys-C is often used because in theory it only cleaves proteins on the C-terminal side of lysine residues, meaning, there will be fewer cleavages, yielding larger product peptides. With middle-down, peptides after the digestion process usually contain more possible protonation sites, which favours MS/MS processes (discussed above).

1.3.4 Proteomics Using Liquid Chromatography-Mass Spectrometry

One of the advantages of MS-based experiments is the ability to use hyphenated techniques, *e.g.* liquid chromatography (LC), gas chromatography (GC), etc. A single proteomic study involves thousands and thousands of proteins, making direct analysis extremely difficult. By coupling a separation technique, *e.g.* LC prior to the MS analysis, also known as LC-MS, it “simplifies” the complexity of the protein/peptide mixtures, making the analysis more feasible.

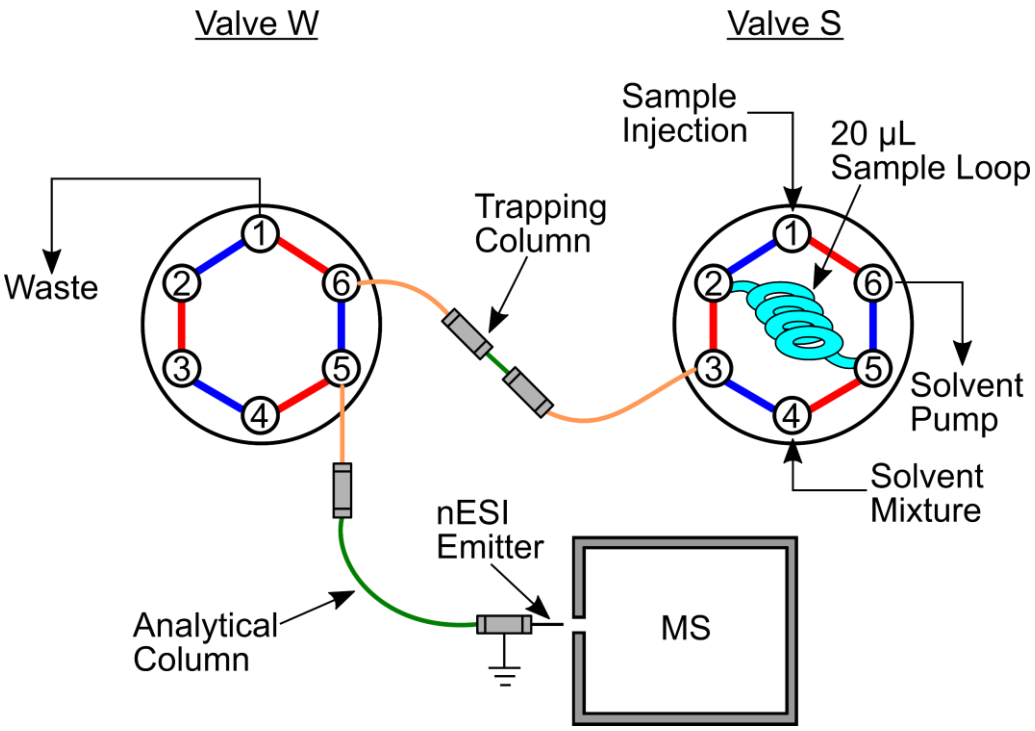


Figure 1.40 Schematic drawing of a LC-MS experiment.

Different valves can be used in a LC set-up. Figure 1.40 shows a common set up using a 6-port valve. Each position is connected to the two neighbouring positions (*e.g.* position 1 connecting with positions 2 and 6) *via* two different channels (indicated with blue and red lines). Only one channel can be used at a time while the other will be blocked. While using a 20 µL sample loop, usually a maximum of 18 µL of sample, are drawn by the solvent pump into the sample loop through the blue channel (1→2→5→6). After that, the rotor rotates (activating the red channel), water from position 4 pushes the sample into the trapping column (shown in green) *via* silica-fused capillary (shown in orange), 4→5→2→3. The trapping column has a wider dimension, allowing a higher flow rate to shorten the trapping time. Packing material (same as those using in the analytical column, *e.g.* C18) that packed in a trapping column which can also act as a desalting purpose by retaining the proteins/peptides and washing away the salt content. Trapping columns are generally very short (~ 2-3 cm) to avoid peak broadening due to separation at this stage. Depending on the sample, if there is a lot of salt present, samples can be trapped in the trapping column and washed for a longer time. Salts are washed out from position 1 of valve W *via* the red channel. After trapping, the valve

rotors of valve S and valve W rotate and activate the blue channel, allowing a shorter distance for the solvent mixture to push the sample through the trapping column into the analytical column (~ 20 cm) *via* a silica-fused capillary. The complex proteomics sample can be separated in the analytical column according to the type of column is employed (*e.g.* reversed phase (RP) column separates the mixture mainly due to the hydrophobicity of the species). Separated samples can then finally be ionised by ESI/nESI and sprayed towards the MS inlet *via* an emitter. Ions generated can also be manipulated and fragmented with the techniques mentioned above.

1.3.5 Separation Methods in LC-MS

Stationary and mobile phases are two important elements in LC. Stationary phase is an immobile material inside a column, which analyte molecules adsorbed onto according to their properties, *e.g.* hydrophobicity, charge etc. Mobile phase refers to the solvents flow through the stationary phase of the column, eluting the analytes by reducing the interaction between the adsorbed molecules and the stationary phase.

1.3.5.1 Separation According to Hydrophobicity of Analytes

Hydrophobicity values of different amino acids have been compiled by Fauchere in 1983 (Table 1.1).[108] The hydrophobicity values range from -1.01 (arginine, most hydrophilic) to 2.25 (tryptophan, most hydrophobic). Proteins and peptides are made up with a combination of different amino acids, therefore, each protein and peptide will have a different hydrophobicity value, which can be calculated by the sum of the corresponding hydrophobicity values of each amino acid involved.[108]

Side Chain of	Hydrophobicity (π_r)	Side Chain of	Hydrophobicity (π_r)
Alanine (A)	0.31	Leucine (L)	1.70
Arginine (R)	-1.01	Lysine (K)	-0.99
Asparagine (N)	-0.60	Methionine (M)	1.23
Aspartic Acid (D)	-0.77	Phenylalanine (F)	1.79
Cysteine (C)	1.54	Proline (P)	0.72
Glutamine (Q)	-0.22	Serine (S)	-0.04
Glutamic Acid (E)	-0.64	Threonine (T)	0.26

Glycine (G)	0	Tryptophan (W)	2.25
Histidine (H)	0.13	Tyrosine (Y)	0.96
Isoleucine (I)	1.80	Valine (V)	1.22

Table 1.1 Hydrophobicity values of the 20 naturally occurring amino acids.[108]

Reversed phase (RP) chromatography is the most common separation method used in LC studies based on the hydrophobicity of the proteins/peptides. The stationary phase of RP columns consists of alkyl chains (C_4/C_8/C_{18}). Different length of the alkyl chains favours the separation of different samples. The shorter alkyl chains are more suitable for larger analytes, therefore C_4 RP columns are commonly used for intact protein separation, while longer alkyl chains are more suitable for smaller analytes, so C_{18} RP columns are commonly used for small molecule/peptide separation. The alkyl chains on an RP column provide a hydrophobic interaction surface for the incoming analyte molecules. Hydrophobic analytes are adsorbed strongly on the surface while the hydrophilic analytes travel through the column without much adsorption (Figure 1.41, a). The red spheres represent highly hydrophobic analyte molecules, which adsorbed strongly with the stationary phase of the RP column and eluted at last, while the green spheres represent less hydrophobic analytes and adsorbed less strongly to the alkyl chains. The blue spheres represent the least hydrophobic analyte molecules, which shows no interaction with the stationary phase and eluted immediately without any retention. The more hydrophobic an analyte is, the higher concentration of organic solvent is required to elute the analyte from the column. RP is the most suitable LC column to couple with MS as the buffers used in RP are salt-free, therefore the LC-MS studies presented in this thesis were carried out using RP columns.

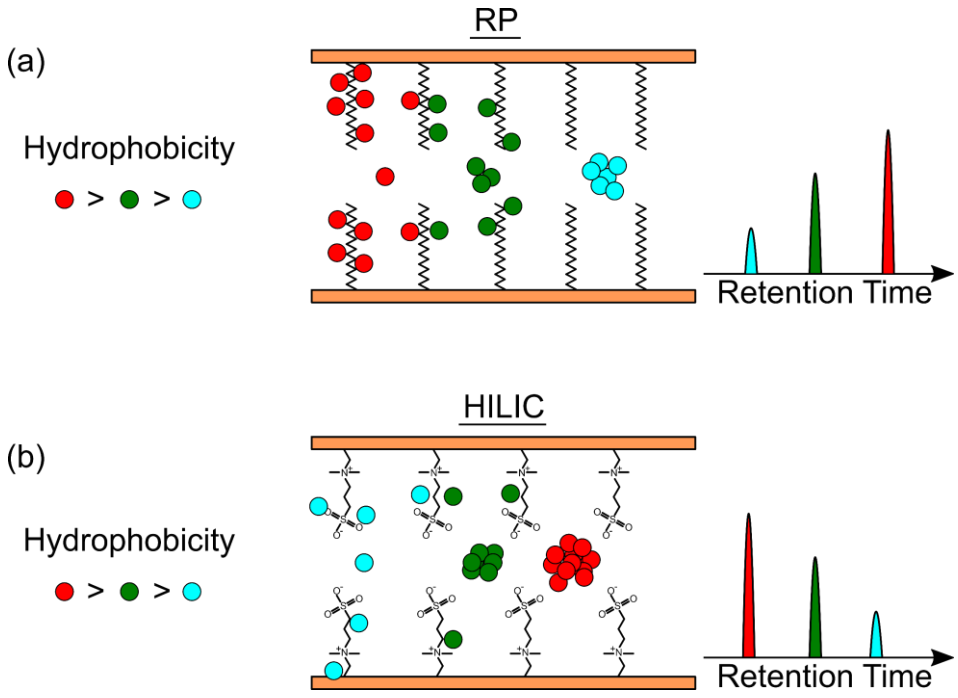


Figure 1.41 Schematic drawing of the two most commonly employed separation techniques based on the hydrophobicity of analytes. a) reversed phase column, b) hydrophilic interaction liquid chromatography column.

Hydrophilic interaction liquid chromatography (HILIC) is an alternative method to separate analyte molecules according to their hydrophobicity. One of the most commonly used stationary phases in HILIC contains zwitterionic sulfobetaine ligands. Unlike RP columns, the stationary phase of HILIC column provides a hydrophilic interaction surface for the incoming analyte molecules. Hydrophilic analytes adsorb strongly to the stationary phase while hydrophobic analyte molecules show little to no interaction to the zwitterionic stationary phase. Therefore, hydrophilic molecules are retained for longer on the column and eluted with a later retention time. HILIC has been shown to achieve a better separation on protein/peptide containing fragile PTMs, such as phospho- or glyco- peptides or proteins.[109,110] Acetonitrile is often used as a loading buffer for HILIC separation while water/methanol is used as an elution buffer due to their high polarity. Ammonium formate ($HCOONH_4$) buffer solution is often used to control the pH to maintain the analytes at a single ionic form for better separation.

1.3.5.2 Separation According to Charge of Analytes

Under different solvent conditions, proteins and peptides present with different charge according to the pH of the environment. Proteins and peptides can be separated based on the polarity in LC.

Strong cation exchange (SCX) relies on the interaction between the analytes and the negatively charged stationary phase (sulfonate, $-SO_3^-$) on the column. Positively charged species interact and retain on the column more effectively while negatively charged species and neutral molecules travel through the column with little to no interaction and elute before the positive species. The higher the magnitude of positive charge, the stronger the electrostatic interaction and the longer it takes for the species to be eluted from the column. Strong anion exchange (SAX) is the opposite to SCX, by having positively charged stationary phase (quaternary amine, $-N(CH_3)_3^+$), negatively charged species are attracted and retained on the column while the positively charged species and neutral molecules are eluted with limited retention. Extremely high salt concentration, *e.g.* 350 mM of KCl or NaCl, is often used to elute the analytes from the column, therefore extra purification steps are needed prior to MS analysis. The schematic representations of SCX and SAX are shown below (Figure 1.42).

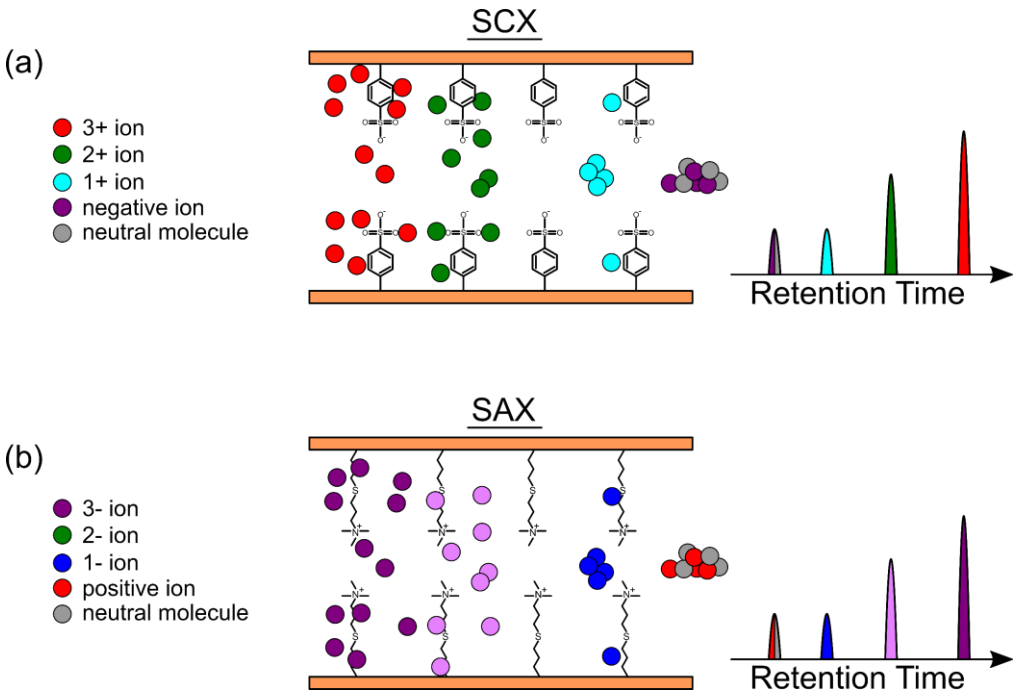


Figure 1.42 Schematic representations of a) strong cation exchange (SCX), and b) strong anion exchange (SAX).

1.3.6 Quantification in Proteomics

A main aspect of proteomics research is to identify and quantify the concentration ratios of protein expressed in cells or tissues in different states/conditions, *e.g.* healthy cell and cancerous cells.[111] The ability to quantify the concentrations/levels of proteins prevent provides an extra dimension of information and this information can often be critical to drug development studies.[112] There are different quantification methods developed by researchers and the three most commonly used techniques are discussed below.

1.3.6.1 Quantification Using Label-Free Analysis

Label-free quantification is one of the most commonly used methods to achieve a relative abundance of proteins/peptides between two samples. Assumptions are taken that under well controlled conditions with sufficient data redundancy, identical peptides across different LC-MS/MS experiments can be compared directly.[112] Sample

preparation and data acquisition usually have to be performed very carefully to minimise external influences on the observed protein quantity. Samples to be compared and quantified should be prepared using identical reagents by a single researcher, to eliminate variations induced be different bottles of reagents and human errors. All analysis should also be performed in a single batch to minimise intra-analysis variations. To quantify proteins/peptides across samples, an internal control is usually added to the sample at known concentrations to evaluate the reproducibility of the experiment.[112] There are different ways to perform a label-free quantification experiment, and they are discussed below.

Ion Abundance

Number of ions can be reflected by the height or the peak area of a peak at a particular *m/z*. Quantity can be measured by directly comparing the ratio of ion abundances between two identical peaks (Figure 1.43, i). However, variations induced by difference in experimental set up will not be accounted for.

Ion Abundance Ratios

Quantification can also be performed by measuring the ratio between the extracted ion chromatogram (EIC) peak area of the ion of interest and the EIC peak area of the internal control in sample 1 and compare to that in sample 2, variations induced between the experimental set up or ionisation efficiency can be accounted for (Figure 1.43, ii). The assumption is made that the internal control is not reacting with the sample and the detection of one does not affect detection of the other and vice versa.

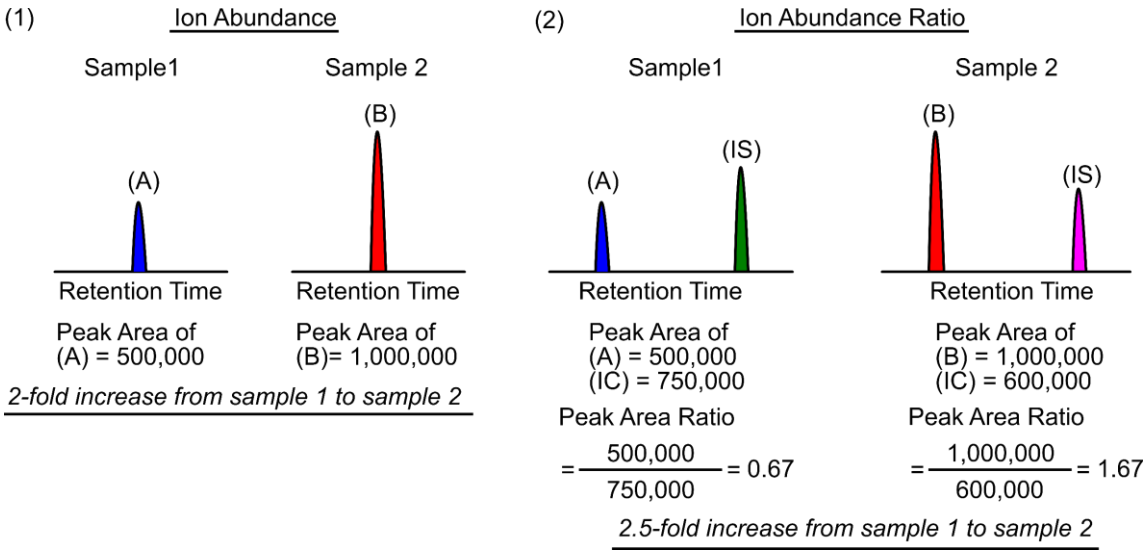


Figure 1.43 Representations of the two label-free quantification methods. i) Ion abundance, ii) Ion abundance ratio. (A) is the peak of interest in sample 1, (B) is the peak of interest in sample 2, and (IS) is the internal standard induced in the corresponding sample.

Spectral Counting

The spectral count for a protein/peptide refers to the number of MS/MS spectra acquired during a LC-MS/MS experiment. Assumption being the more abundant a protein/peptide is, the more likely it is will be selected for MS/MS analysis. However, this method is also limited to runs where no active inclusion can be used for the MS/MS, meaning MS/MS of the same peptide will be acquired repeatedly, lowering the possible number of protein identifications. The total spectral count should be the same across samples under identical data acquisition and can be represented by the equation below:

$$S_{1n} = \frac{\sum S_{2i}}{\sum S_{1i}} \qquad \text{[Eq. 21]}$$

where s is the spectral count for species n, 1 and 2 are the sample numbers and i is the i^{th} species.

1.3.6.2 Quantification Using iTRAQ/TMT

Isobaric tag for relative and absolute quantification (iTRAQ) and tandem mass tag (TMT) have also been commonly used for quantification in proteomics studies.[111]

Quantification is performed by tagging a cleavable molecule in each sample onto the peptides. The chemical tag is made up by three components, a reporter group, a balance group, and an amine specific peptide reactive group (N-hydroxysuccinimide, NHS) (Figure 1.44, a). Different reporter groups with varying isotope encoding, usually with a substitution of a ^{12}C for a ^{13}C (neutromer),[113,114] are paired up with their corresponding balancing groups, therefore, all of the tags will be exactly the same in mass (Figure 1.44, b). These chemical tags attach to all peptides in a protein digest *via* free amines at the peptide N-terminus and on the side chain of lysine residues.[115] Different samples with different tags induced into will be mixed together. The peptide of interest from all samples will appear as a single species because masses of all tags are the same. Upon fragmentation, *e.g.* CAD, the reporter group will be dissociated, and their intensities is directly proportional to the peptide quantities (Figure 1.44, c).

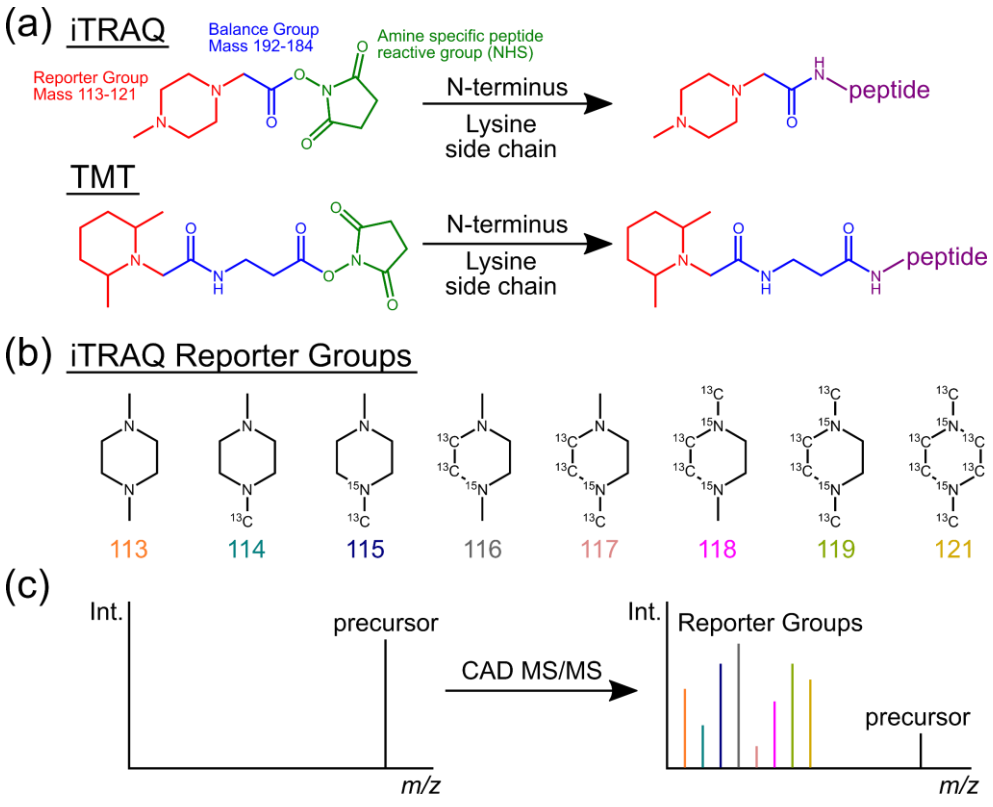


Figure 1.44 Quantification using iTRAQ/TMT. a) Schematic representing of the reaction between the chemical tag and peptides (iTRAQ and TMT), b) chemical structures of iTRAQ reporter groups, c) quantification of peptides by reporter groups upon CAD MS/MS.[111]

1.3.6.3 Quantification Using SILAC

Addition of stable isotope amino acids into the cell culture (SILAC) is another method that commonly used for quantification studies in proteomics. The principle of SILAC is based on inducing stable isotope labelled amino acids, such as ^{13}C or ^{15}N-labelled arginine or lysine, into the entire proteome during protein metabolism.[116] Two populations of cells are grown in two different culture media, "light" or "heavy", where the light medium contains amino acids with natural isotopes, and the heavy medium contains amino acids with heavy isotopes. Proteins from cells grown in the heavy medium will be made of amino acids with heavy isotopes after enough cell division events (also known as the adaption stage),[117] while the proteins from cells grown in the light medium will have natural amino acids. After the adaption stage, in the experimental stage, cells from the two culture mediums will be mixed together, the mixture will often be analysed by LC-MS/MS, the quantification of SILAC is based on calculating the ratio of the heavy isotope labelled peptide and the natural isotope peptide. The change in signal intensities from the light sample to heavy sample allow a quantitative comparison of their relative abundance in the mixture to be made. This method assumes the kinetic isotope effect has a negligible impact on the biological system.

1.4 Post-Translational Modifications

Post-translational modifications (PTMs) refer to the modification of proteins after the protein biosynthesis event. PTMs are covalently bound and induced enzymatically.[118] Proteins are synthesised by ribosomes using translation messenger RNA (mRNA) into polypeptides, which may then undergo a modification process and form mature protein. PTMs play a fundamental role in regulating the folding of proteins and other important cellular functions, such as cell signalling,[119] regulating cell proliferation,[120] etc. There are many different types of PTMs, and PTM discovery is a very popular research topic. One of most commonly studied PTMs, oxidation, by MS is discussed below.

1.4.1 Oxidation

Oxidation is a commonly observed PTM in proteins; consisting of the inclusion of oxygen at groups into species and often induced by reactive oxygen species (ROS). Cancer cells are known to produce more ROS compared to healthy cells,[121] however, an excessive amount of ROS can lead to a redox imbalance (oxidative stress), triggering cell death/apoptosis. Therefore, by studying the location of the oxidation site, this may provide more information about the cancer cells, and developing more effective anti-cancer agents, ideally by selectively inducing oxidative stress in cancerous cells only.[93] However, in order to target the possible oxidised species by MS and include these oxidation modifications in the database searches, it is important to study oxidation of different amino acids first as not all amino acids can be oxidised. It is worth noting that some oxidised products can be used to diagnose the source of oxidation/the identity of the ROS. With the aid of the ultra-high mass accuracy by FT-ICR, the mass difference causing by oxidation can be measured and help localise the oxidation site in the protein/peptide.

Generation of ROS is a commonly used technique in PACT and PDT (discussed above) to induce oxidative stress near or in cancerous cells. ROS is often produced after the energy transfer from the excited sensitiser to molecular oxygen (Figure 1.45), to yield singlet oxygen ($^1O_2$) and oxidises the surrounding targets (*e.g.* proteins).[122]

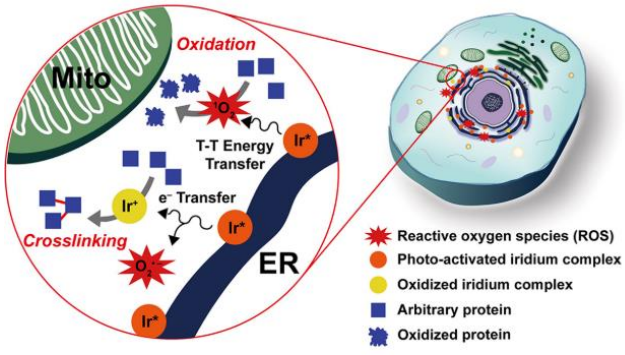


Figure 1.45 Schematic representation of protein modifications in the mitochondria (Mito) and endoplasmic reticulum (ER) by a photoactivable Ir(III) complex. (Reproduced from Nam *et al.* 2016 with permission).[96]

References

(1) Aebersold, R.; Mann, M. Mass Spectrometry-Based Proteomics. *Nature* **2003**, *422* (6928), 198–207.

(2) Dettmer, K., Aronov, P.A., Hammock, B. D. Mass Spectrometry-Based Metabolomics. *Mass Spectrom. Rev.* **2007**, *26*, 51–78.

(3) Rodgers, R.P., Marshall, A. G. *Petroleomics: Advanced Characterization of Petroleum-Derived Materials by Fourier Transform Ion Cyclotron Resonance Mass Spectrometry (FT-ICR MS)*; 2007, Springer, New York, USA.

(4) Hoffmann, E. De; Stroobant, V. *Mass Spectrometry Principles and Applications Third Edition*; 2007, Wiley, Hoboken, New Jersey, USA.

(5) Dempster, A. J. A New Method of Positive Ray Analysis. *Phys. Rev.* **1918**, *11*, 316–325.

(6) Dongre, V. G.; Karmuse, P. P.; Nimbalkar, M. M.; Singh, D.; Kumar, A. Application of GC-EI-MS for the Identification and Investigation of Positional Isomer in Primaquine, an Antimalarial Drug. *J. Pharm. Biomed. Anal.* **2005**, *39* (1–2), 111–116.

(7) Li, Y.; Zhang, H.; Hu, J.; Xue, F.; Li, Y.; Sun, C. A GC-EI-MS-MS Method for Simultaneous Determination of Seven Adulterants in Slimming Functional Foods. *J. Chromatogr. Sci.* **2012**, *50* (10), 928–933.

(8) Nikolaev, E. Victor Talrose: An Appreciation. *J. Mass Spectrom.* **1998**, *33* (6), 499–501.

(9) Munson, M. S. B.; Field, F. H. Chemical Ionization Mass Spectrometry. I. General Introduction. *J. Am. Chem. Soc.* **1966**, *88* (12), 2621–2630.

(10) Field, F. H.; Munson, M. S. B. Chemical Ionization Mass Spectrometry. V. Cycloparaffins. *J. Am. Chem. Soc.* **1967**, *89* (17), 4272–4280.

(11) Rayleigh, Lord. XX. On the Equilibrium of Liquid Conducting Masses Charged with Electricity. *London, Edinburgh, Dublin Philos. Mag. J. Sci.* **1882**, *14* (87),

184–186.

(12) Smith, J. N.; Flagan, R. C.; Beauchamp, J. L. Droplet Evaporation and Discharge Dynamics in Electrospray Ionization. *J. Phys. Chem. A* **2002**, *106* (42), 9957–9967.

(13) Konermann, L.; Ahadi, E.; Rodriguez, A. D.; Vahidi, S. Unraveling the Mechanism of Electrospray Ionization. *Anal. Chem.* **2013**, *85* (1), 2–9.

(14) Wilm, M.; Mann, M. Analytical Properties of the Nanoelectrospray Ion Source. *Anal. Chem.* **1996**, *68* (1), 1–8.

(15) Karas, M.; Bachmann, D.; Bahr, U.; Hillenkamp, F. Matrix-Assisted Ultraviolet Laser Desorption of Non-Volatile Compounds. *Int. J. Mass Spectrom. Ion Process.* **1987**, *78* (C), 53–68.

(16) Karas, M.; Bachmann, D.; Hillenkamp, F. Influence of the Wavelength in High-Irradiance Ultraviolet Laser Desorption Mass Spectrometry of Organic Molecules. *Anal. Chem.* **1985**, *57* (14), 2935–2939.

(17) Tanaka, K.; Waki, H.; Ido, Y.; Akita, S.; Yoshida, Y. Protein and Polymer Analyses up to m/z 100 000 by Laser Ionization Time-of-Flight Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **1988**, *2* (8), 151–153.

(18) Wolfgang, P.; Steinwedel, H. Notizen: Ein Neues Massenspektrometer Ohne Magnetfeld. *Zeitschrift für Naturforsch. A* **1953**, *8* (7), 448–450.

(19) March, R. E.; McMahon, A. W.; Londry, F. A.; Alfred, R. L.; Todd, J. F. J.; Vedel, F. Resonance Excitation of Ions Stored in a Quadrupole Ion Trap. Part 1. A Simulation Study. *Int. J. Mass Spectrom. Ion Process.* **1989**, *95* (2), 119–156.

(20) Hopfgartner, G.; Varesio, E.; Tschäppät, V.; Grivet, C.; Bourgogne, E.; Leuthold, L. A. Triple Quadrupole Linear Ion Trap Mass Spectrometer for the Analysis of Small Molecules and Macromolecules. *J. Mass Spectrom.* **2004**, *39* (8), 845–855.

(21) Wolff, M. M.; Stephens, W. E. A Pulsed Mass Spectrometer with Time Dispersion. *Rev. Sci. Instrum.* **1953**, *24* (8), 616–617.

(22) Mamyrin, B. A.; Karataev, V. I.; Shmikk, D. V.; Zagulin, V. A. The Mass-Reflectron, a New Nonmagnetic Time-of-Flight Mass Spectrometer with High Resolution. *Sov. Phys. - JETP* **1973**, *37* (1), 45–48.

(23) Comisarow, M. B.; Marshall, A. G. Fourier Transform Ion Cyclotron Resonance Spectroscopy. *Chem. Phys. Lett.* **1974**, *25* (2), 282–283.

(24) Comisarow, M. B. B. M. B.; Marshall, A. G. G. A. G. Frequency-Sweep Fourier Transform Ion Cyclotron Resonance Spectroscopy. *Chem. Phys. Lett.* **1974**, *26* (4), 489–490.

(25) Walker, L. R.; Tfaily, M. M.; Shaw, J. B.; Hess, N. J.; Paša-Tolić, L.; Koppenaal, D. W. Unambiguous Identification and Discovery of Bacterial Siderophores by Direct Injection 21 Tesla Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Metallomics* **2017**, *9* (1), 82–92.

(26) Wootton, C. A.; Sanchez-Cano, C.; Lopez-Clavijo, A. F.; Shaili, E.; Barrow, M. P.; Sadler, P. J.; O'Connor, P. B. Sequence-Dependent Attack on Peptides by Photoactivated Platinum Anticancer Complexes. *Chem. Sci.* **2018**, *9* (10), 2733–2739.

(27) Lam, Y. P. Y.; Wootton, C. A.; Hands-Portman, I.; Wei, J.; Chiu, C. K. C.; Romero-Canelon, I.; Lermyte, F.; Barrow, M. P.; O'Connor, P. B. Does Deamidation of Islet Amyloid Polypeptide Accelerate Amyloid Fibril Formation? *Chem. Commun.* **2018**, *54* (98), 13853–13856.

(28) Marshall, A. G.; Hendrickson, C. L. Fourier Transform Ion Cyclotron Resonance Detection: Principles and Experimental Configurations. *Int. J. Mass Spectrom.* **2002**, *215* (1–3), 59–75.

(29) Marshall, A. G.; Hendrickson, C. L.; Jackson, G. S. Fourier Transform Ion Cyclotron Resonance Mass Spectrometry: A Primer. *Mass Spectrom. Rev.* **1998**, *17* (1), 1–35.

(30) Amster, I. J. Fourier Transform Mass Spectrometry. *J Mass Spectrom* **1996**, *31* (September), 1325–1337.

(31) Marshall, A. G.; Roe, D. C. Theory of Fourier Transform Ion Cyclotron Resonance Mass Spectroscopy: Response to Frequency-Sweep Excitation. *J. Chem. Phys.* **1980**, *73* (4), 1581–1590.

(32) Marshall, A. G.; Wang, T. C. L.; Ricca, T. L. Tailored Excitation for Fourier Transform Ion Cyclotron Mass Spectrometry. *J. Am. Chem. Soc.* **1985**, *107* (26), 7893–7897.

(33) Comisarow, M. B. Signal Modeling for Ion Cyclotron Resonance. *J. Chem. Phys.* **1978**, *69* (9), 4097–4104.

(34) Marshall, A. G. Milestones in Fourier Transform Ion Cyclotron Resonance Mass Spectrometry Technique Development. *Int. J. Mass Spectrom.* **2000**, *200* (1–3), 331–356.

(35) Caravatti, P.; Allemann, M. The 'Infinity Cell': A New Trapped-ion Cell with Radiofrequency Covered Trapping Electrodes for Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Org. Mass Spectrom.* **1991**, *26* (5), 514–518.

(36) Beu, S. C.; Laude, D. A. Elimination of Axial Ejection during Excitation with a Capacitively Coupled Open Trapped-Ion Cell for Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Anal. Chem.* **1992**, *64* (2), 177–180.

(37) Scigelova, M.; Hornshaw, M.; Giannakopulos, A.; Makarov, A. Fourier Transform Mass Spectrometry. *Mol. Cell. Proteomics* **2018**, *10* (7), 1–19.

(38) Jennings, K. R. Colllision-Induced Decompositions of Aromatic Ions. *Int. J. Mass Spectrom. Ion Phys.* **1968**, *1* (3), 227–235.

(39) Little, D. P.; Speir, J. P.; Senko, M. W.; O'Connor, P. B.; McLafferty, F. W. Infrared Multiphoton Dissociation of Large Multiply Charged Ions for Biomolecule Sequencing. *Anal. Chem.* **1994**, *66* (18), 2809–2815.

(40) Ly, T.; Julian, R. R. Ultraviolet Photodissociation: Developments towards Applications for Mass-Spectrometry-Based Proteomics. *Angew. Chemie - Int. Ed.* **2009**, *48* (39), 7130–7137.

(41) Zubarev, R.; Kelleher, N. L.; McLafferty, F. W. Electron Capture Dissociation of Multiply Charged Protein Cations. A Nonergodic Process. *J. Am. Chem. Soc* **1998**, *120* (16), 3265–3266.

(42) Kalli, A.; Grigorean, G.; Håkansson, K. Electron Induced Dissociation of Singly Deprotonated Peptides. *J. Am. Soc. Mass Spectrom.* **2011**, *22* (12), 2209–2221.

(43) Wootton, C. A.; Sanchez-Cano, C.; Liu, H.-K.; Barrow, M. P.; Sadler, P. J.; O'Connor, P. B. Binding of an Organo-Osmium(II) Anticancer Complex to Guanine and Cytosine on DNA Revealed by Electron-Based Dissociations in High Resolution Top-Down FT-ICR Mass Spectrometry. *Dalton Trans.* **2015**, *44* (8), 3624–3632.

(44) Syka, J. E. P.; Coon, J. J.; Schroeder, M. J.; Shabanowitz, J.; Hunt, D. F. Peptide and Protein Sequence Analysis by Electron Transfer Dissociation Mass Spectrometry. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101* (26), 9528–9533.

(45) Roepstorff, P. Proposal for a Common Nomenclature for Sequence Ions in Mass Spectra of Peptides. *Biomed. Mass Spectrom.* **1984**, *11* (11), 601.

(46) Cooper, H. J.; Hakansson, K.; Marshall, A. G. The Role of Electron Capture Dissociation in Biomolecular Analysis. *Mass Spectrom. Rev.* **2005**, *24* (2), 201–222.

(47) Li, H.; Lin, T. Y.; Van Orden, S. L.; Zhao, Y.; Barrow, M. P.; Pizarro, A. M.; Qi, Y.; Sadler, P. J.; O'Connor, P. B. Use of Top-down and Bottom-up Fourier Transform Ion Cyclotron Resonance Mass Spectrometry for Mapping Calmodulin Sites Modified by Platinum Anticancer Drugs. *Anal. Chem.* **2011**, *83* (24), 9507–9515.

(48) Xu, F.; Xu, Q.; Dong, X.; Guy, M.; Guner, H.; Hacker, T. A.; Ge, Y. Top-down High-Resolution Electron Capture Dissociation Mass Spectrometry for Comprehensive Characterization of Post-Translational Modifications in Rhesus Monkey Cardiac Troponin I. *Int. J. Mass Spectrom.* **2011**, *305* (2–3), 95–102.

(49) Zhang, X.; Li, H.; Moore, B.; Wongkongkathep, P.; Ogorzalek Loo, R. R.; Loo,

J. A.; Julian, R. R. Radical-Directed Dissociation of Peptides and Proteins by Infrared Multiphoton Dissociation and Sustained off-Resonance Irradiation Collision-Induced Dissociation with Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **2014**, *28* (24), 2729–2734.

(50) Breuker, K.; Oh, H. B.; Lin, C.; Carpenter, B. K.; McLafferty, F. W. Nonergodic and Conformational Control of the Electron Capture Dissociation of Protein Cations. *Proc. Natl. Acad. Sci.* **2004**, *101* (39), 14011–14016.

(51) Zhurov, K. O.; Fornelli, L.; Wodrich, M. D.; Laskay, A.; Tsybin, Y. O. Principles of Electron Capture and Transfer Dissociation Mass Spectrometry Applied to Peptide and Protein Structure Analysis. *Chem. Soc. Rev. Chem. Soc. Rev* **2013**, *42* (42), 5014–5030.

(52) Sawicka, A.; Skurski, P.; Hudgins, R. R.; Simons, J. Model Calculations Relevant to Disulfide Bond Cleavage *via* Electron Capture Influenced by Positively Charged Groups. *J. Phys. Chem. B* **2003**, *107* (48), 13505–13511.

(53) Syrstad, E. A.; Tureček, F. Toward a General Mechanism of Electron Capture Dissociation. *J. Am. Soc. Mass Spectrom.* **2005**, *16* (2), 208–224.

(54) Li, H.; O'Connor, P. B. Electron Capture Dissociation of Disulfide, Sulfur-Selenium, and Diselenide Bound Peptides. *J. Am. Soc. Mass Spectrom.* **2012**, *23* (11), 2001–2010.

(55) Zubarev, R. A.; Horn, D. M.; Fridriksson, E. K.; Kelleher, N. L.; Kruger, N. A.; Lewis, M. A.; Carpenter, B. K.; McLafferty, F. W. Electron Capture Dissociation for Structural Characterization of Multiply Charged Protein Cations. *Anal. Chem.* **2000**, *72* (3), 563–573.

(56) Mikhailov, V. A.; Cooper, H. J. Activated Ion Electron Capture Dissociation (AI ECD) of Proteins: Synchronization of Infrared and Electron Irradiation with Ion Magnetron Motion. *J. Am. Soc. Mass Spectrom.* **2009**, *20* (5), 763–771.

(57) Ge, Y.; Lawhorn, B. G.; ElNaggar, M.; Strauss, E.; Park, J. H.; Begley, T. P.; McLafferty, F. W. Top down Characterization of Larger Proteins (45 KDa) by

Electron Capture Dissociation Mass Spectrometry. *J. Am. Chem. Soc.* **2002**, *124* (4), 672–678.

(58) Shou, W.; Verma, R.; Annan, R. S.; Huddleston, M. J.; Chen, S. L.; Carr, S. A.; Deshaies, R. J. Mapping Phosphorylation Sites in Proteins by Mass Spectrometry. *Methods Enzymol.* **2002**, *351*, 279–296.

(59) Chamot-Rooke, J.; van der Rest, G.; Dalleu, A.; Bay, S.; Lemoine, J. The Combination of Electron Capture Dissociation and Fixed Charge Derivatization Increases Sequence Coverage for O-Glycosylated and O-Phosphorylated Peptides. *J. Am. Soc. Mass Spectrom.* **2007**, *18* (8), 1405–1413.

(60) Wootton, C. A.; Millett, A. J.; Lopez-Clavijo, A. F.; Chiu, C. K. C.; Barrow, M. P.; Clarkson, G. J.; Sadler, P. J.; O'Connor, P. B. Structural Analysis of Peptides Modified with Organo-Iridium Complexes, Opportunities from Multi-Mode Fragmentation. *Analyst* **2019**, *144*, 1575–1581.

(61) Haselmann, K. F.; Budnik, B. A.; Zubarev, R. A. Electron Capture Dissociation of B2+ Peptide Fragments Reveals the Presence of the Acylium Ion Structure. *Rapid Commun. Mass Spectrom.* **2000**, *14* (23), 2242–2246.

(62) Budnik, B. A.; Haselmann, K. F.; Zubarev, R. A. Electron Detachment Dissociation of Peptide Di-Anions: An Electron-Hole Recombination Phenomenon. *Chem. Phys. Lett.* **2001**, *342* (3–4), 299–302.

(63) Narayanan, D. L.; Saladi, R. N.; Fox, J. L. Ultraviolet Radiation and Skin Cancer. *Int. J. Dermatol.* **2010**, *49*, 978–986.

(64) Shannon, A. M.; Bouchier-Hayes, D. J.; Condron, C. M.; Toomey, D. Tumour Hypoxia, Chemotherapeutic Resistance and Hypoxia-Related Therapies. *Cancer Treat. Rev.* **2003**, *29* (4), 297–307.

(65) Circu, M. L.; Aw, T. Y. Reactive Oxygen Species, Cellular Redox Systems, and Apoptosis. *Free Radic. Biol. Med.* **2010**, *48* (6), 749–762.

(66) Moreno, P. M. D.; Pego, A. P. Therapeutic Antisense Oligonucleotides against Cancer: Hurdling to the Clinic. *Front. Chem.* **2014**, *2* (October), 1–7.

(67) Bednarski, P.; Mackay, F.; Sadler, P. Photoactivatable Platinum Complexes. *Anticancer. Agents Med. Chem.* **2008**, *7* (1), 75–93.

(68) Aykul, S.; Martinez-Hackert, E. Determination of Half-Maximal Inhibitory Concentration Using Biosensor-Based Protein Interaction Analysis. *Anal. Biochem.* **2016**, *508*, 97–103.

(69) Yang, J.; Kopeček, J. Macromolecular Therapeutics. *J. Control. Release* **2014**, *190*, 288–303.

(70) Reedijk, J. Platinum Anticancer Coordination Compounds: Study of DNA Binding Inspires New Drug Design. *Eur. J. Inorg. Chem.* **2009**, No. 10, 1303–1312.

(71) Ndagi, U.; Mhlongo, N.; Soliman, M. E. Metal Complexes in Cancer Therapy – An Update from Drug Design Perspective. *Drug Des. Devel. Ther.* **2017**, *11*, 599–616.

(72) Pizarro, A. M.; Sadler, P. J. Unusual DNA Binding Modes for Metal Anticancer Complexes. *Biochimie* **2009**, *91* (10), 1198–1211.

(73) Sava, G.; Bergamo, A.; Dyson, P. J. Metal-Based Antitumour Drugs in the Post-Genomic Era: What Comes Next? *Dalton Trans.* **2011**, *40* (36), 9069–9075.

(74) Wong, E.; Giandornenico, C. M. Current Status of Platinum-Based Antitumor Drugs. *Chem. Rev.* **1999**, *99* (9), 2451–2466.

(75) Hearn, J. M.; Romero-Canelón, I.; Qamar, B.; Liu, Z.; Hands-Portman, I.; Sadler, P. J. Organometallic Iridium(III) Anticancer Complexes with New Mechanisms of Action: NCI-60 Screening, Mitochondrial Targeting, and Apoptosis. *ACS Chem. Biol.* **2013**, *8* (6), 1335–1343.

(76) Ivanov, A. I.; Christodoulou, J.; Parkinson, J. A.; Barnham, K. J.; Tucker, A.; Woodrow, J.; Sadler, P. J. Cisplatin Binding Sites on Human Albumin. *J. Biol. Chem.* **1998**, *273* (24), 14721–14730.

(77) Pillaire, M. J.; Hoffmann, J. S.; Defais, M.; Villani, G. Replication of DNA

Containing Cisplatin Lesions and Its Mutagenic Consequences. *Biochimie* **1995**, *77* (10), 803–807.

(78) Alderden, R. A.; Hall, M. D.; Hambley, T. W. The Discovery and Development of Cisplatin. *J. Chem. Educ.* **2009**, *83* (5), 728.

(79) Zayed, A.; Shoeib, T.; Taylor, S. E.; Jones, G. D. D.; Thomas, A. L.; Wood, J. P.; Reid, H. J.; Sharp, B. L. Determination of Pt-DNA Adducts and the Sub-Cellular Distribution of Pt in Human Cancer Cell Lines and the Leukocytes of Cancer Patients, Following Mono- or Combination Treatments, by Inductively-Coupled Plasma Mass Spectrometry. *Int. J. Mass Spectrom.* **2011**, *307* (1–3), 70–78.

(80) Centerwall, C. R.; Tacka, K. A.; Kerwood, D. J.; Goodisman, J.; Toms, B. B.; Dubowy, R. L.; Dabrowiak, J. C. Modification and Uptake of a Cisplatin Carbonato Complex by Jurkat Cells. *Mol. Pharmacol.* **2006**, *70* (1), 348–355.

(81) Hearn, J. M.; Romero-Canelón, I.; Munro, A. F.; Fu, Y.; Pizarro, A. M.; Garnett, M. J.; McDermott, U.; Carragher, N. O.; Sadler, P. J. Potent Organo-Osmium Compound Shifts Metabolism in Epithelial Ovarian Cancer Cells. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112* (29), E3800–E3805.

(82) Millett, A. J.; Habtemariam, A.; Romero-Canelón, I.; Clarkson, G. J.; Sadler, P. J. Contrasting Anticancer Activity of Half-Sandwich Iridium(III) Complexes Bearing Functionally Diverse 2-Phenylpyridine Ligands. *Organometallics* **2015**, *34* (11), 2683–2694.

(83) Peacock, A. F. A.; Habtemariam, A.; Fernández, R.; Walland, V.; Fabbiani, F. P. A.; Parsons, S.; Aird, R. E.; Jodrell, D. I.; Sadler, P. J. Tuning the Reactivity of Osmium(II) and Ruthenium(II) Arene Complexes under Physiological Conditions. *J. Am. Chem. Soc.* **2006**, *128* (5), 1739–1748.

(84) Morris, R. E.; Aird, R. E.; Del Socorro Murdoch, P.; Chen, H.; Cummings, J.; Hughes, N. D.; Parsons, S.; Parkin, A.; Boyd, G.; Jodrell, D. I.; et al. Inhibition of Cancer Cell Growth by Ruthenium(II) Arene Complexes. *J. Med. Chem.* **2001**, *44* (22), 3616–3621.

(85) Ang, W. H.; Casini, A.; Sava, G.; Dyson, P. J. Organometallic Ruthenium-Based Antitumor Compounds with Novel Modes of Action. *J. Organomet. Chem.* **2011**, *696* (5), 989–998.

(86) Barragán, F.; López-Senín, P.; Salassa, L.; Betanzos-Lara, S.; Habtemariam, A.; Moreno, V.; Sadler, P. J.; Marchán, V. Photocontrolled DNA Binding of a Receptor-Targeted Organometallic Ruthenium(II) Complex. *J. Am. Chem. Soc.* **2011**, *133* (35), 14098–14108.

(87) Lovejoy, K. S.; Lippard, S. J. Non-Traditional Platinum Compounds for Improved Accumulation, Oral Bioavailability, and Tumor Targeting. *Dalton Trans.* **2009**, 10651–10659.

(88) Wang, J.; Liu, G.; Leung, K.; Loffroy, R.; Lu, P.-X.; J. Wáng, Y. Opportunities and Challenges of Fluorescent Carbon Dots in Translational Optical Imaging. *Curr. Pharm. Desi*. 2015, pp 5401–5416.

(89) Shi, H.; Romero-Canelón, I.; Hreusova, M.; Novakova, O.; Venkatesh, V.; Habtemariam, A.; Clarkson, G. J.; Song, J. I.; Brabec, V.; Sadler, P. J. Photoactivatable Cell-Selective Dinuclear Trans-Diazidoplatinum(IV) Anticancer Prodrugs. *Inorg. Chem.* **2018**, *57* (22), 14409–14420.

(90) Farrer, N. J.; Salassa, L.; Sadler, P. J. Photoactivated Chemotherapy (PACT): The Potential of Excited-State d-Block Metals in Medicine. *Dalton Trans.* **2009**, No. 48, 10690–10701.

(91) Shi, H.; Clarkson, G. J.; Sadler, P. J. Dual Action Photosensitive Platinum(II) Anticancer Prodrugs with Photoreleasable Azide Ligands. *Inorganica Chim. Acta* **2019**, *489* (December 2018), 230–235.

(92) Davies, M. J. The Oxidative Environment and Protein Damage. *Biochim. Biophys. Acta - Proteins Proteomics* **2005**, *1703* (2), 93–109.

(93) Zhang, P.; Chiu, C. K. C.; Huang, H.; Lam, Y. P. Y.; Habtemariam, A.; Malcomson, T.; Paterson, M. J.; Clarkson, G. J.; O'Connor, P. B.; Chao, H.; et al. Organoiridium Photosensitizers Induce Specific Oxidative Attack on Proteins

within Cancer Cells. *Angew. Chemie Int. Ed.* **2017**, 14898–14902.

(94) Gracanin, M.; Hawkins, C. L.; Pattison, D. I.; Davies, M. J. Singlet-Oxygen-Mediated Amino Acid and Protein Oxidation: Formation of Tryptophan Peroxides and Decomposition Products. *Free Radic. Biol. Med.* **2009**, *47* (1), 92–102.

(95) Josefsen, L. B.; Boyle, R. W. Photodynamic Therapy and the Development of Metal-Based Photosensitisers. *Met. Based. Drugs* **2008**, *2008*, 1–23.

(96) Nam, J. S.; Kang, M. G.; Kang, J.; Park, S. Y.; Lee, S. J. C.; Kim, H. T.; Seo, J. K.; Kwon, O. H.; Lim, M. H.; Rhee, H. W.; et al. Endoplasmic Reticulum-Localized Iridium(III) Complexes as Efficient Photodynamic Therapy Agents *via* Protein Modifications. *J. Am. Chem. Soc.* **2016**, *138* (34), 10968–10977.

(97) Wasinger, V. C.; Cordwell, S. J.; Poljak, A.; Yan, J. X.; Gooley, A. A.; Wilkins, M. R.; Duncan, M. W.; Harris, R.; Williams, K. L.; Humphery-Smith, I. Progress with Gene-product Mapping of the Mollicutes: Mycoplasma Genitalium. *Electrophoresis* **1995**, *16* (1), 1090–1094.

(98) Walther, T. C.; Mann, M. Mass Spectrometry-Based Proteomics in Cell Biology. *J. Cell Biol.* **2010**, *190* (4), 491–500.

(99) McDonald, W. H.; Yates, J. R. Shotgun Proteomics and Biomarker Discovery. *Dis. Markers* **2002**, *18* (2), 99–105.

(100) Gillette, M. A.; Carr, S. A. Quantitative Analysis of Peptides and Proteins in Biomedicine by Targeted Mass Spectrometry. *Nat. Methods* **2013**, *10* (1), 28–34.

(101) Anderson, L. C.; Dehart, C. J.; Kaiser, N. K.; Fellers, R. T.; Smith, D. F.; Greer, J. B.; Leduc, R. D.; Blakney, G. T.; Thomas, P. M.; Kelleher, N. L.; et al. Identification and Characterization of Human Proteoforms by Top-Down LC-21 Tesla FT-ICR Mass Spectrometry. *J. Proteome Res.* **2017**, *16* (2), 1087–1096.

(102) Elzek, M. A.; Rodland, K. D. Proteomics of Ovarian Cancer: Functional Insights and Clinical Applications. *Cancer Metastasis Rev.* **2015**, *34* (1), 83–96.

(103) Wootton, C. A.; Lam, Y. P. Y.; Willetts, M.; van Agthoven, M. A.; Barrow, M. P.; Sadler, P. J.; O′Connor, P. B. Automatic Assignment of Metal-Containing Peptides in Proteomic LC-MS and MS/MS Data Sets. *Analyst* **2017**, *142* (11), 2029–2037.

(104) Fornelli, L.; Toby, T. K.; Schachner, L. F.; Doubleday, P. F.; Srzenti??, K.; DeHart, C. J.; Kelleher, N. L. Top-down Proteomics: Where We Are, Where We Are Going? *J. Proteomics* **2016**, 10–11.

(105) Chen, B.; Brown, K. A.; Lin, Z.; Ge, Y. Top-Down Proteomics: Ready for Prime Time? *Anal. Chem.* **2018**, *90* (1), 110–127.

(106) Ahn, J.; Cao, M. J.; Yu, Y. Q.; Engen, J. R. Accessing the Reproducibility and Specificity of Pepsin and Other Aspartic Proteases. *Biochim. Biophys. Acta - Proteins Proteomics* **2013**, *1834* (6), 1222–1229.

(107) Suttapitugsakul, S.; Xiao, H.; Smeekens, J.; Wu, R. Evaluation and Optimization of Reduction and Alkylation Methods to Maximize Peptide Identification with MS-Based Proteomics. *Mol. Biosyst.* **2017**, *13* (12), 2574–2582.

(108) Fauchere, J. L.; Pliska, V. Hydrophobic Parameters Pi of Amino Acid Side Chains from the Partitioning of N-Acetyl-Amino Acid Amides. *Eur. J. Med. Chem.* **1983**, *18* (4), 369–375.

(109) Ahn, Y. H.; Kim, J. Y.; Yoo, J. S. Quantitative Mass Spectrometric Analysis of Glycoproteins Combined with Enrichment Methods. *Mass Spectrosc. Rev.* **2015**, *34*, 148–165.

(110) Boersema, P. J.; Mohammed, S.; Heck, A. J. R. Hydrophilic Interaction Liquid Chromatography (HILIC) in Proteomics. *Anal. Bioanal. Chem.* **2008**, *391* (1), 151–159.

(111) Wiese, S.; Reidegeld, K. A.; Meyer, H. E.; Warscheid, B. Protein Labeling by ITRAQ: A New Tool for Quantitative Mass Spectrometry in Proteome Research. *Proteomics* **2007**, *7* (3), 340–350.

(112) Wong, J. W. H.; G., C. *An Overview of Label-Free Quantification Methods in*

Proteomics by Mass Spectrometry; 2009, Humana Press, New Jersey, USA.

(113) Naylor, B. C.; Porter, M. T.; Wilson, E.; Herring, A.; Lofthouse, S.; Hannemann, A.; Piccolo, S. R.; Rockwood, A. L.; Price, J. C. DeuteRater: A Tool for Quantifying Peptide Isotope Precision and Kinetic Proteomics. *Bioinformatics* **2017**, *33* (10), 1514–1520.

(114) Smith, R.; Taylor, R. M.; Prince, J. T. Current Controlled Vocabularies Are Insufficient to Uniquely Map Molecular Entities to Mass Spectrometry Signal. *BMC Bioinformatics* **2015**, *16* (7), S2.

(115) Unwin, R. D. Quantification of Proteins by ITRAQ. In *LC-MS/MS in Proteomics - Methods and Applications*; Cutillas, P. R., Timms, J. F., Eds.; Humana Press, 2010; pp 205–215.

(116) Chen, X.; Wei, S.; Ji, Y.; Guo, X.; Yang, F. Quantitative Proteomics Using SILAC: Principles, Applications, and Developments. *Proteomics* **2015**, *15* (18), 3175–3192.

(117) Mann, M. Functional and Quantitative Proteomics Using SILAC. *Nat. Rev. Mol. Cell Biol.* **2006**, *7* (12), 952–958.

(118) Mann, M.; Jensen, O. N. Proteomic Analysis of Post-Translational Modifications. *Nat. Biotechnol.* **2003**, *21* (March), 255–261.

(119) del Monte, F.; Agnetti, G. Protein Post-Translational Modifications and Misfolding: New Concepts in Heart Failure. *Proteomics - Clin. Appl.* **2014**, *8* (7–8), 534–542.

(120) Seo, J.-W.; Lee, K.-J. Post-Translational Modifications and Their Biological Functions: Proteomic Analysis and Systematic Approaches. *J. Biochem. Mol. Biol.* **2011**, *37* (1), 35–44.

(121) Szatrowski, T. P.; Nathan, C. F. Production of Large Amounts of Hydrogen Peroxide by Human Tumor Cells. *Cancer Res.* **1991**, *51* (3), 794–798.

(122) Oattison, D. I.; Rahmanto, A. S.; Davis, M. J. Photo-Oxidation of Proteins.

Photochem. Photobiol. Sci. **2012**, *11*, 38.

Chapter 2

Metallodrug-Peptide Interactions Studied by Ultra-High Resolution Mass Spectrometry

Together with the increased use of platinum-based anticancer agents over the past decades, there is also a significant increase in platinum resistant cancer cases. Transition metal-based complexes possess a range of geometries and activities against many different forms of cancer. Potent piano-stool compounds, such as those based on ruthenium, iridium, and osmium, have been shown to be extremely effective and selective against cancerous cells, including resistant strains.[1]

Bottom-up MS analysis is often employed by researchers due to the relatively low instrument requirement compared to top-down approach. However, sensitive modification such as metallodrug binding can present challenges to the experiments. Herein, microwave digestion, an alternative digestion method was investigated to compare to standard tryptic digestion, looking into possibilities of achieving an effective digestion in 15 min rather than the standard 16 h tryptic digestion. Different fragmentation and hyphenated techniques were also used to investigate the effectiveness of studying system with fragile modifications. The osmium complex (Os1-Cl) was synthesised by Dr. Carlos Sanchez-Cano. Sample preparation and all data were produced, analysed, and interpreted by the author of this thesis.

A manuscript entitled “Metallodrug-Peptide Interactions Studied by Ultra-High Resolution Mass Spectrometry” by Cookson K. C. Chiu, Yuko P. Y. Lam, Christopher A. Wootton, Mark P. Barrow, Peter J. Sadler, and Peter B. O’Connor, has been prepared for submission to the Journal of the American Society for Mass Spectrometry (JASMS) based directly on the contents and findings of this chapter.

Abstract

Here we investigate the effectiveness of alternative digestion methods for bottom up analysis, and develop a methodology for localising metallodrug-binding sites. Bovine serum albumin (BSA) has long been used for bottom-up digestion and LC-MS analysis optimisation as it produces a wide variety of representative tryptic peptides upon digestion, therefore it was used as a model protein to obtain a mixture of peptides for bottom-up metallo-interaction analysis. The methodology developed herein identified metallated species in a drug-peptide mixture effectively, and same methodology can potentially be applied to more complex biomolecule systems, such as drug-cell MS analysis, to evaluate their behaviour towards commonly encountered biomolecules.

Introduction

Metal-based chemotherapies are at the forefront of anticancer therapy, with predominantly platinum based compounds, such as cisplatin, carboplatin, and oxaliplatin being used in chemotherapy treatments for various types of cancer.[3] However, the number of platinum-resistant cancer cases has continuously increased with the increased use of platinum-based anti-cancer agents.[4] Therefore, potently cytotoxic and anti-proliferative compounds have been developed using a range of other transition metals, including ruthenium,[5–8] iridium,[9–12] osmium,[2,13–15], rhodium,[16], gold,[17,18] as well as new platinum-based designs.[19–22] These new transition metal-based drugs have been shown to have varying mechanisms of action (MoA) against cancerous cell components, while some are able to disrupt vital processes in cells in order to kill cancer cells, such as interaction between the metallodrug and mitochondria.[23,24] It is therefore of much interest to investigate the ability of these new metallodrugs to react with different biomolecules, to further study different possible MoA against cancerous cells. Metallodrugs are often designed as pro-drugs (non-active compounds), and ideally only activated when the drug is inside/near the cancerous cells, to increase effectiveness and selectivity towards cancerous cells, minimising the side effects from reacting with healthy cells inside the body. During the journey towards the target, the composition of metallodrugs is often found to change, from drug activation (desired), or exposure to a

change in environment; pH, temperature, salt concentration, etc., (often undesired). Therefore, it is important to characterise the behaviour as well as composition of the metallodrug, to provide more information on how the metallodrug interacts with the biomolecules.

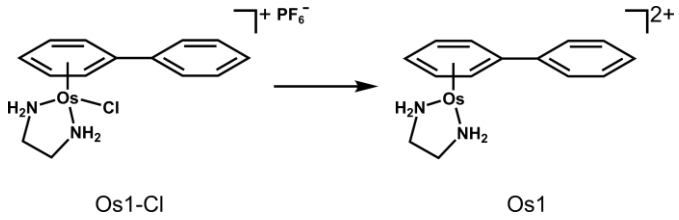
Many analytical techniques have been utilised to study metal-biomolecule interactions, often ^{1}H NMR,[25] or x-ray crystallography.[26] These techniques possess powerful ability to provide structural information and sometimes able to identify and locate the modifications on biomolecules. However, a mixture of products often introduces issues to these analytical techniques due to the inherent lack of separation ability. Therefore, to address these issues, other analytical techniques are being explored which possess the ability of handling mixtures of adducts effectively. Mass spectrometry (MS) is increasingly being employed to study biological systems, due to its high sensitivity and ability to cope with complex samples. With the aid of tandem mass spectrometry (MS/MS), fragmentation of the ionised target species provides an in-depth chemical and structural information of the complex biological system. An extensive array of fragmentation techniques has been developed, including collisionally activated dissociation (CAD),[27] electron capture dissociation (ECD),[28] infrared multi-photon dissociation (IRMPD),[29] etc. Different fragmentation techniques have their own advantages, allowing a more comprehensive analysis.

Fourier transform-ion cyclotron resonance mass spectrometry (FT-ICR MS) uses electric and magnetic fields to trap ions of interest within a Penning trap during MS/MS experiments and detection.[30,31] Latest commercial setups allow resolving powers of 500,000 – 10,000,000 or above,[32] which is not achievable by other mass spectrometers. With the capability of isolating ions of interest using a quadrupole, applying collisional activation, and electron/photon-based dissociation techniques upon ions of interests, FT-ICR MS provides the largest array of fragmentation techniques and highest mass accuracy of any mass spectrometer, with mass errors into sub parts-per-million (ppm) and even parts-per-billion (ppb) ranges.

The interaction between an osmium-(II) complex and a DNA 12-mer was studied by Wootton *et al.* The chlorido ligand was found to be dissociated and produced

a doubly charged species (Os1, Scheme 2.1). The DNA-12mer was observed to be osmiated (osmium bound) and the binding locations were found to be guanine$_3$ and cytosine$_{10}$ with the aid of electron detachment dissociation MS/MS.[2]

Here we investigated the effectiveness of microwave digestion compared to standard tryptic digestion of BSA. A bottom-up metallo-interaction methodology was also developed to study the interaction between model BSA tryptic digest and an osmium-(II) arene anticancer complex, Os1-Cl (Scheme 2.1). Ultra-high resolution mass spectrometry together with various MS/MS techniques to investigate the binding location of the metallodrug on the biomolecule. Reversed phase (RP) nano-liquid chromatography (nLC) was also coupled to FT-ICR MS to separate the complex generated from the tryptic digestion of BSA, minimising the ion suppression/space charge effects from having too many ions in the ICR cell during excitation and detection, allowing better observation and study on the osmiated species.



Scheme 2.1: The Os(II) arene complex, Os1-Cl studied in this work and the species (Os1) detected *via* nano-electrospray ionisation mass spectrometry (nESI-MS).

Experimental

Materials

Bovine serum albumin (BSA, ≥ 98%, fatty acid free), formic acid, ammonium bicarbonate (AmmBic), dithiothreitol (DTT), iodoacetamide (IAA), and trypsin were all purchased from Sigma Aldrich (Dorset, England). LCMS grade acetonitrile (MeCN) was purchased from VWR (Lutterworth, UK). Ultra-pure water (18.2 $M\Omega cm^{-1}$) was obtained from a Milli-Q UV III water purification system (Milli-Q, Hertfordshire, UK)

Preparation of osmium complex

$[(\eta^6\text{-bip})Os(en)Cl]PF_6$ ($Os1\text{-}Cl\cdot PF_6$) was synthesised and characterised as described previously by Dr. Carlos Sanchez-Cano.[33]

Preparation of BSA tryptic digest

Standard Digestion

BSA was reduced using 50 mM of DTT, alkylated with 100 mM of IAA, and tryptic digested with 1 µg/µL of trypsin in 100 mM of AmmBic at 37 ^oC for 16 h. Digested samples were then desalted using C18 SPE cartridges (Thermo Fisher Scientific, Waltham, UK) as follows; washed twice with 1% formic acid-acidified water and eluted using 80% MeCN/aqueous solution. Desalted samples were then filtered through 3 kDa molecular weight cut-off filters (Millipore, Hertfordshire, UK) to remove any digested proteins above 3 kDa. The resulting BSA digest samples had a concentration of 1 mg/mL, with an assumption of a 100% digestion efficiency from initial protein concentration.

Microwave Digestion

BSA was reduced using 50 mM of DTT, and separated into two batches, one batch alkylated with 100 mM of IAA, and the other one without. Then 12 µL of 1 µg/µL of trypsin was added into the sample. The sample mixture was microwave-digested using a CEM Discover SP-D 80 Automated Microwave Digestion System (Buckingham, UK). Digestion temperature was optimised to 100 ^oC, ramp time of 1 min, hold time of 15 min, and power of 20 W to achieve optimum digestion. The

resulting BSA digest sample had a concentration of 1 mg/mL total peptide/protein content.

Preparation of Os1-BSA digest reaction

Osmium complex, Os1-Cl (250 μM) was reacted with the BSA digest, to a 1:1 molar ratio using 1 mg/mL protein digest and an average peptide mass of 1500 Da. The reaction mixture was incubated under 37 ^{o}C for one day. Samples for direct infusion MS were diluted 100-fold with 50:50 (v/v) H_2O:MeCN and 0.1% formic acid. Samples for nLC-FT-ICR MS analysis were diluted 25-fold with water.

Nano-LC Separation

All LC separations were achieved using an EASY-nano-LC II system (Proxeon, Hemel Hempstead, UK) with a 3 cm, C18 RP trapping column (150 μm, 3 μm particle size, pore size of 300 Å), and a 25 cm, C18 RP analytical column (75 μm, 3 μm particle size, pore size of 300 Å). Separation of tryptic digest samples was achieved using an acidified water/MeCN (solvent + 0.1% formic acid) gradient from 5% B (MeCN) to 30% B over 30 min before a 30 min wash of 80% B. 12 μL (7 pmol) of sample was injected for each nLC-MS or nLC-MS/MS run.

Mass Spectrometry

All samples were analysed using nano-electrospray ionisation (nESI) for increased sensitivity and lower the sample consumption compared to standard electrospray ionisation (ESI). All experiments were carried out on a SolariX FT-ICR mass spectrometer, fitted with a 12 tesla actively shielded superconducting magnet (Bruker Daltonik GmbH, Bremen, Germany). All samples were sprayed in positive ion mode. For direct infusion MS experiments, ions were accumulated for 0.5 s in the hexapole-based collision cell before transferred to the infinity cell for detection.[34]

For CAD MS/MS experiments, ions of interest were isolated using the front-end quadrupole using an isolation window of 10 *m/z*. Selected ions were then accelerated into argon collision gas at 12 V (for unmodified peptides) and 5 V for (osmiated species). Ions were continuously accumulated for 4 s before transmission and detection.

For ECD MS/MS experiments, ions of interest were isolated using the front-end quadrupole using an isolation window of 10 *m/z*, then accumulated in the hexapole for 4 s. Isolated ions of interest were then transferred and trapped in the infinity cell. Trapped ions were then irradiated for 0.15 s with 1.2 eV electrons produced from a 1.5 A hollow cathode dispenser *via* an extraction lens held at 18 V.

All nLC-MS spectra were peak picked with the Sophisticated Numerical Annotation Procedure – Liquid Chromatography (SNAP-LC) algorithm developed by Wootton *et al.*[35] Osmiated-species were peak picked using the following parameters: quality factor of 0.95, a signal-to-noise ratio (S/N) threshold of 2, and a maximum charge state of 5+, a constant building block (base elementary composition) of C = 4.9384, N = 1.3577, O = 1.4773, S = 0.0417, H = 7.7583.

Auto MS/MS with a prefer list of osmiated species picked out by the SNAP-LC algorithm was performed to fragment the selected osmiated species. Intensity threshold of $1x10^6$ was used and all singly-charged ions are excluded from the fragmentation list. The isolated osmium species were further accumulated for 3 s in the hexapole before being transferred into the infinity cell for ECD MS/MS.

All spectra were internally calibrated using a quadratic calibration function,[36] manually interpreted and assigned *via* DataAnalysis v4.2 (Bruker Daltonik GmbH, Bremen, Germany).

Results and Discussion

Serum albumin is the most abundant protein in blood plasma, and it is responsible for drug transportation in bodies, therefore BSA was chosen as a model protein for the method development of metallodrug binding site studies. BSA digestion was achieved using two different methods, standard digestion and microwave digestion. Microwave digestion was achieved in 15 min while standard digestion often requires 16 hr or more to achieve an effective level of digestion. BSA was reduced with 50 mM of DTT and separated into two batches of sample, one then treated with IAA to achieve alkylation of cysteine and the other one without. Both batches then underwent the same microwave digestion process and yielded different results in observed peptides.

Peptides observed in the non-alkylated sample (Figure 2.1a) were larger in size and more highly charged compared to the ones observed in the alkylated sample (Figure 2.1b). One of advantages of using electrospray ionisation (ESI) is the ability to generate multiply charged molecular ions from the analytes without fragmenting them. In FT-ICR MS, the detected signal is proportional to the charge of the species too, therefore, it is more sensitive to multiply-charged ions compared to singly-charged ions. Figure 1b shows a wide range of singly-charged ions, believed to be due to the extra vibrational energy transferred into BSA by the microwave, and overcome the bond strength between residues, therefore, some very small peptides (3-residue short peptides are not often observed) were produced. While in Figure 2.1a, doubly-charged or even triply-charged peptides are observed, which is believed that the protein structure is still "protected" by the reforming disulfide bonds, by not using an alkylating reagent to block off the free thiols from DTT reduction. Therefore, the only location that can carry a charge is the N-terminus of the short peptide, while in the non-alkylated sample, there are more sites on the peptide can carry charges, *e.g.* histidine (H), lysine (K), arginine (R), and the N-terminus.

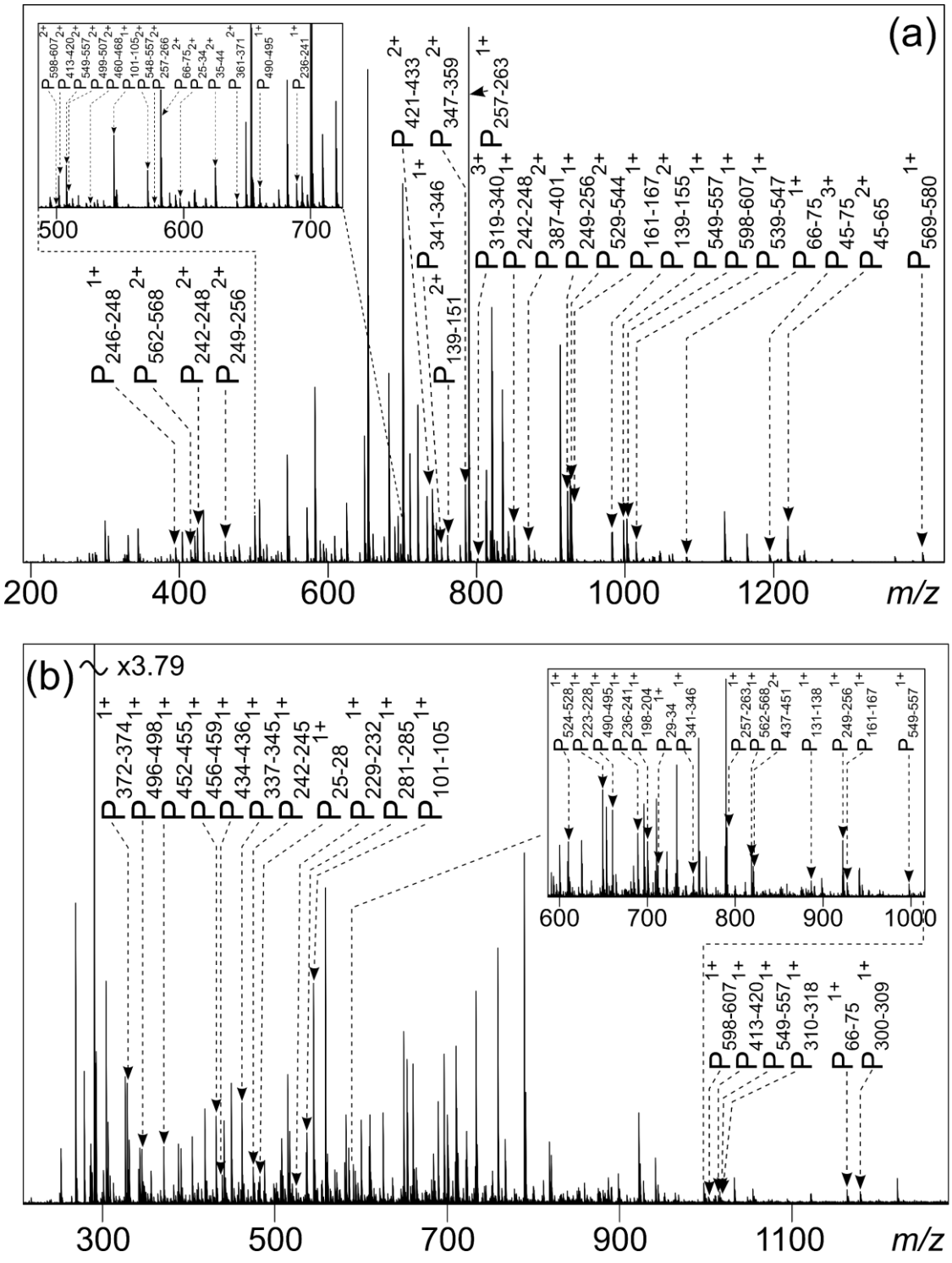


Figure 2.1 Microwave digestion of BSA. a) reduced with DTT b) reduced with DTT and alkylated with IAA. Observed peptides (P) are annotated with their amino acid sequence numbers.

BSA was tryptic-digested overnight to create a mixture of peptides and reacted with the Os1-Cl in a 1:1 ratio for a day at 37 °C. The sample was diluted and analysed by FT-ICR MS *via* direct infusion (Figure 2.2). The resulting mass spectrum was internally calibrated by using unmodified BSA peptides (Table S2.1). SNAP peak picking algorithm was used to look for species containing an osmium isotope pattern (Figure 2.2 inset). Four osmiated species were detected using SNAP (labelled in purple). It is important to identify the peptide, but also to locate exactly which residue the osmium complex is binding to, and the type of interaction, such as covalent interaction or non-covalent interaction. Different tandem mass spectrometry (MS/MS) techniques were used to investigate the binding.

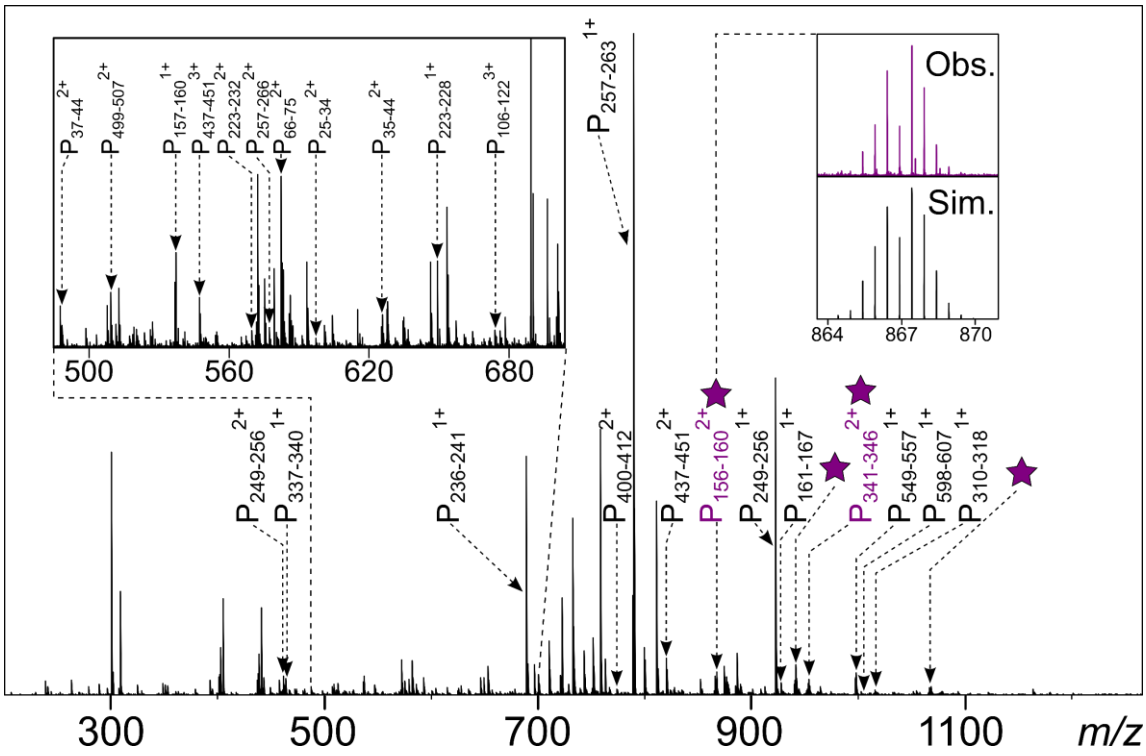


Figure 2.2 Full mass spectrum of the [Os1-Cl]-BSA digest reaction mixture. Purple stars and purple assignments represent osmiated species. Inset on the right shows the comparison between the observed species (top) and the simulation (bottom).

Collisionally activated dissociation (CAD) was chosen to perform the MS/MS experiment on the osmiated species ($[P_{156\text{-}160}]^{2+}$). The dissociation process is performed

inside the hexapole collision cell, located after the quadrupole. The species of interest (867.42385 *m/z*) was isolated inside the quadrupole and accelerated into a cloud of inert argon gas molecules in the hexapole collision cell. CAD is a slow heating process which tends to cleave the amide bond on the protein/peptide backbone. However, if there are any weaker interactions, such as non-covalent interactions, π-π stacking, hydrogen bonding, etc., these interactions will often be broken before the protein/peptide backbone. 5 V of collision energy was applied to osmiated species (Figure 2.3, Table S2.2), and unfortunately, the osmium complex dissociated readily, and was observed as a free metallodrug ion (Os1). However, important information was also obtained from the CAD MS/MS spectrum, the peptide (residue 156-160, KFWGK) was observed as a dimer and the interaction between the osmium complex and the peptide dimer was weak.

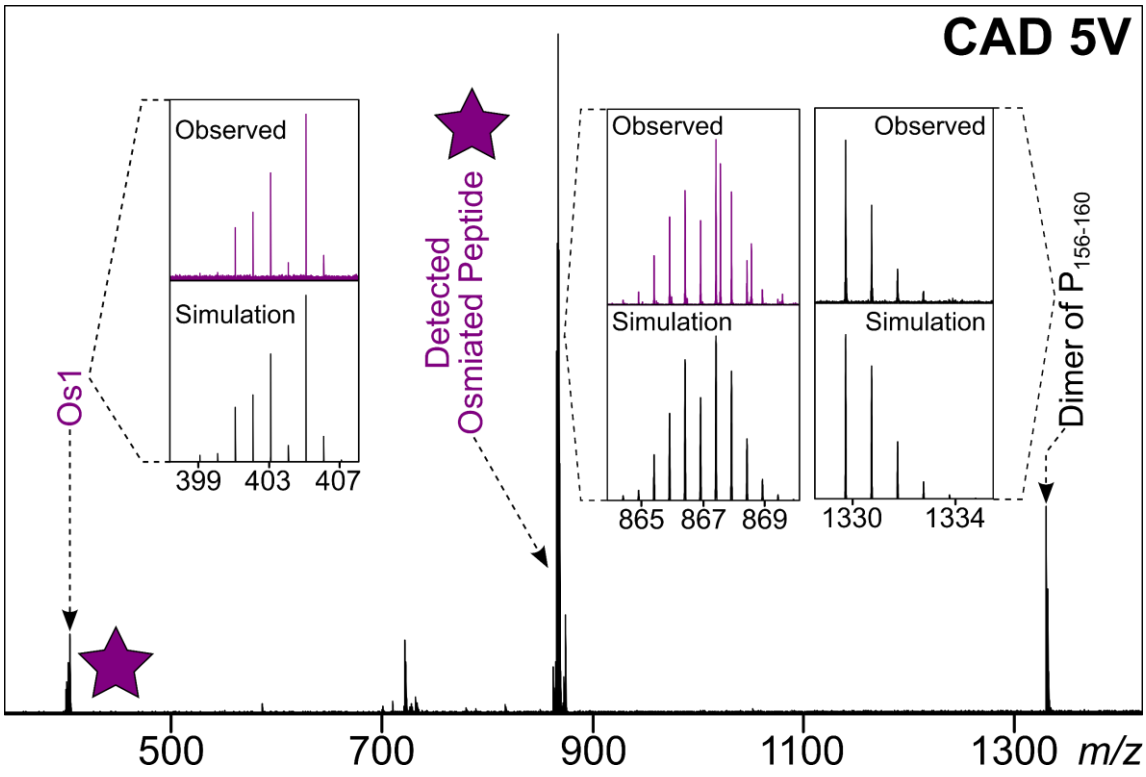


Figure 2.3 CAD MS/MS of the osmiated species (867.42385 *m/z*). Inset shows the comparison between the observed species (top) and the simulation (bottom).

Other fragmentation techniques were required to investigate the binding while the osmium complex is retained. Electron capture dissociation (ECD) is a radical driven

fragmentation process,[28] meaning unlike CAD fragmenting the species using a slow heating approach, ECD cleaves bonds preferentially, such as the N-C_α bond on the protein/peptide backbone, therefore, weak interactions/labile modifications, are normally preserved with ECD MS/MS. Figure 2.4 shows that the osmium complex (Os1) is successfully retained on the peptide. Note that the two positive charges are contributed by the osmium complex by losing the chlorido ligand from Os1-Cl and binding to the neutral peptide, which is often observed in metallated species,[2,10] especially if the metal complex itself is a charge carrier (Table S2.3).

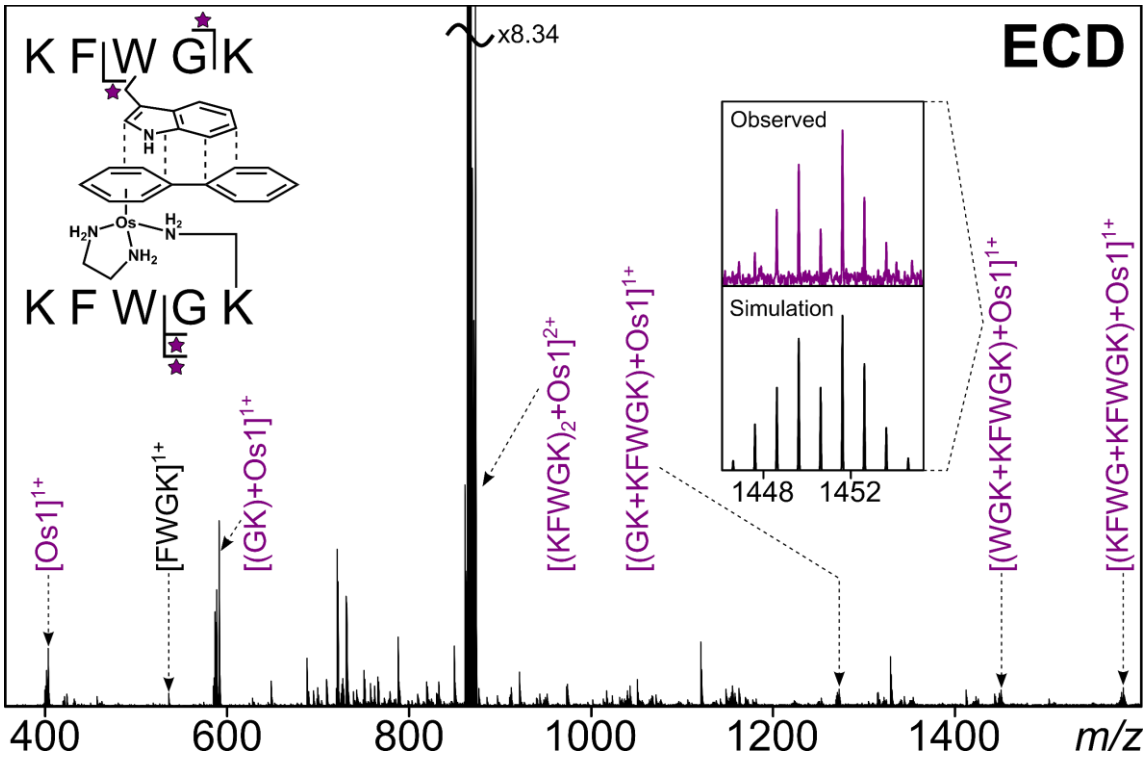


Figure 2.4 ECD MS/MS of the osmiated species (867.42385 *m/z*). Osmium containing species are labelled in purple. Inset shows the comparison between the observed species (top), [(WGK+KFWGK)]+Os1]$^{2+}$, and the simulation (bottom).

By combining information gathered from the low energy (5 V) CAD MS/MS and ECD MS/MS spectra, the conclusion can be made that the binding interaction of Os1 to the dimer peptide was weak, therefore, even applying a low collision energy to the osmiated species caused the dissociation of the metallodrug. From the fragmentation pattern observed in Figure 2.4, the binding interactions in the osmiated

species were suggested to be a coordinate covalent bond, donated from nitrogen on the side chain of a lysine (K) residue and a weak π-π interaction between the phenyl ring of Os1 and the indole side chain of the tryptophan (W) residue. A proposed structure of the osmiated species is shown below in Figure 2.5.

MKWVTFISLL LLFSSAYSRG VFRRDTHKSE IAHRFKDLGE EHFKGLVLIA FSQYLQQCPF DEHVKLVNEL
TEFAKTCVAD ESHAGCEKSL HTLFGDELCK VASLRETYGD MADCCEKQEP ERNECFLSHK DDSPDLPKLK
PDPNTLCDEF KADEKFWGK YLYEIARRHP YFYAPELLYY ANKYNGVFQE CCQAEDKGAC LLPKIETMRE
KVLASSARQR LRCASIQKFG ERALKAWSVA RLSQKFPKAE FVEVTKLVTD LTKVHKECCH GDLLECADDR
ADLAKYICDN QDTISSKLKE CCDKPLLEKS HCIAEVEKDA IPENLPPLTA DFAEDKDVCK NYQEAKDAFL
GSFLYEYSRR HPEYAVSVLL RLAKEYEATL EECCAKDDPH ACYSTVFDKL KHLVDEPQNL IKQNCDQFEK
LGEYGFQNAL IVRYTRKVPQ VSTPTLVEVS RSLGKVGTRC CTKPESERMP CTEDYLSLIL NRLCVLHEKT
PVSEKVTKCC TESLVNRRPC FSALTPDETY VPKAFDEKLF TFHADICTLP DTEKQIKKQT ALVELLKHKP
KATEEQLKTV MENFVAFVDK CCAADDKEAC FAVEGPKLVV STQTALA

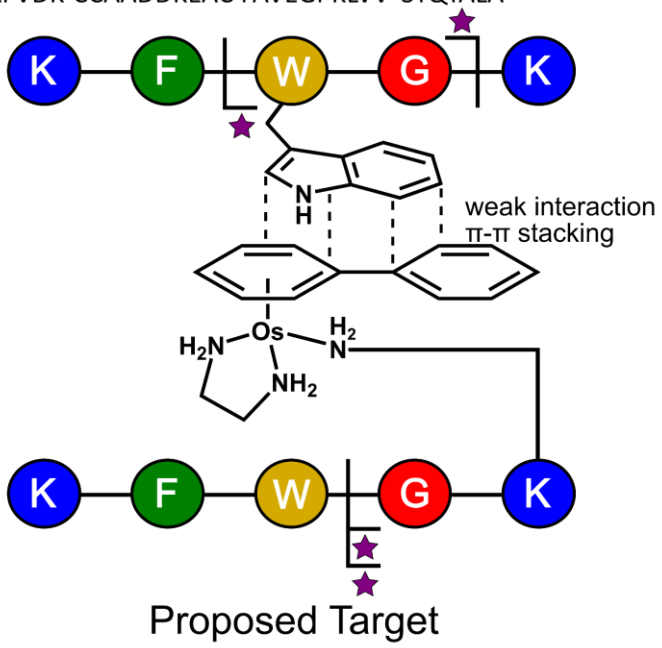


Figure 2.5 Full protein sequence of BSA (top), the species of interest is highlighted in purple. A proposed binding interaction between Os1 and the peptide dimer is shown (bottom right), fragments with purple stars represent the presence of osmium.

By direct infusion of the (Os1-Cl)-BSA reaction products, a few osmiated species were observed. However, they appeared to be in low intensity, which restricted effective isolation and fragmentation. Low intensity can be caused by either inherently low ion abundance in the sample or more often by space charge effects or saturating the

ICR cell with undesired ions.[38] Therefore, online-separation was performed by coupling a nano LC (nLC) to the mass spectrometer, so that the peptide mixture can be separated prior to MS analysis. Peptides were separated using a reversed phase (RP) C18 column according to the hydrophobicity of the peptide.[39] The more hydrophobic the peptide is, the more it is retained on the column and eluted at later retention time. nLC-MS was performed first in order to obtain *m/z* information of the separated peptide mixture throughout the retention times (Figure 2.6a). The nLC-MS data were peak-picked using Sophisticated Numerical Annotation Procedure - Liquid Chromatography (SNAP-LC) algorithm developed by Wootton *et al.*[35], to pick out automatically any osmiated species in each spectrum along the retention times (Figure 2.6b). nLC-ECD MS/MS was then performed using a prefer list of osmiated peaks generated by the algorithm in the previous step (Figure 2.6c). Extraction ion chromatograms (EIC) of osmiated peaks from the prefer list can be obtained from the nLC-MS/MS data, and the binding of Os1 on each detected osmiated peptide investigated (Figure 2.6d).

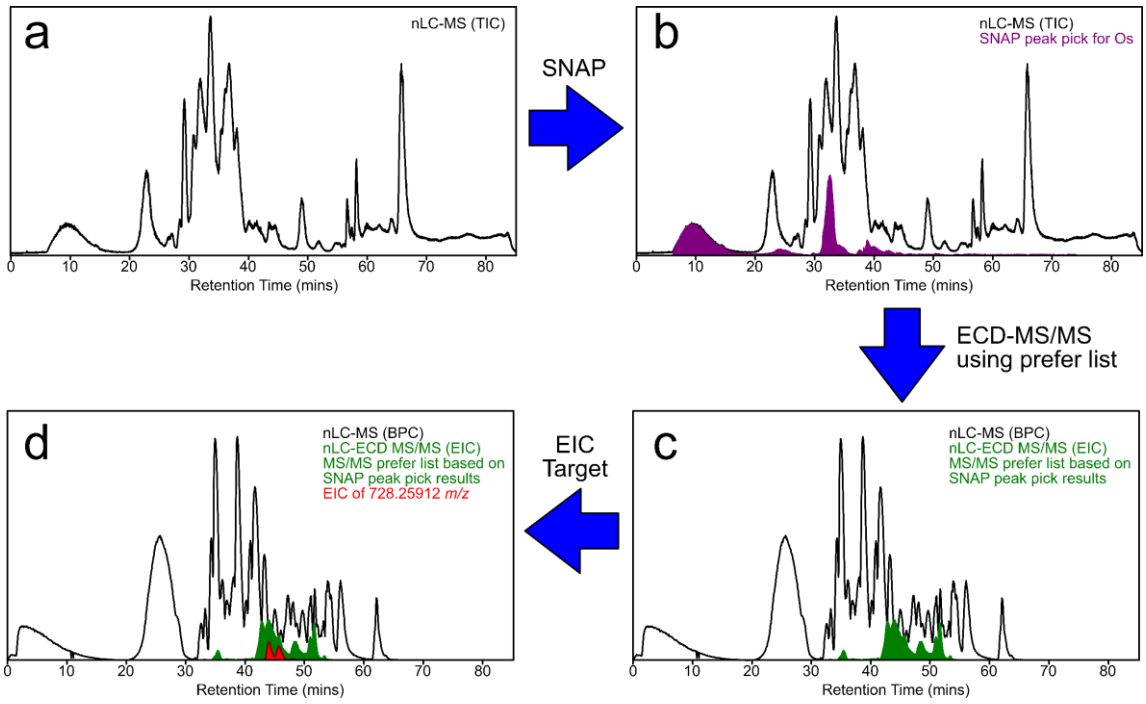


Figure 2.6 nLC-MS/MS experimental workflow to obtain metal-binding information. a) Total ion chromatogram (TIC) of the nLC-MS run is shown, and obtained *m/z* information of the peptide mixture, b) SNAP-LC peak picking algorithm was applied to pick out osmiated species (shown in purple in the chromatogram) and generate a prefer list of osmiated species, c) nLC-ECD MS/MS using the prefer list, EIC of the fragmented osmiated species are shown in green in the chromatogram, d) example of visualising the ECD MS/MS of one of the osmiated peaks (728.25912 *m/z*, shown in red in the chromatogram).

Figure 2.7 shows the ECD MS/MS spectrum of an osmiated species at 728.25912 *m/z* (Table S2.4). It is worth noting that, the unmodified peptide (pENCDQFEK) is observed. This is believed to be due to a dissociation of the osmium drug during transmission with some of the precursor ions or an electron is captured at the metal centre, causing instability of the complex and dissociated the ligand/peptide. Despite the dissociation of the osmium drug, some useful osmiated fragments were still observed, allowing the characterisation of the binding site of the osmium drug. A series of osmiated z fragments were observed, from z_2 to z_5, narrowed down the binding region to the two residues (glutamic acid (E) and lysine (K) near the C-terminus). 728.25912

m/z was observed as a doubly-charged species, and the two positive charges were contributed by the osmium metal centre. The nitrogen from the lysine side chain donates two electrons to form a coordinate covalent bond with the metal centre, as shown in the proposed structure in Figure 2.7.

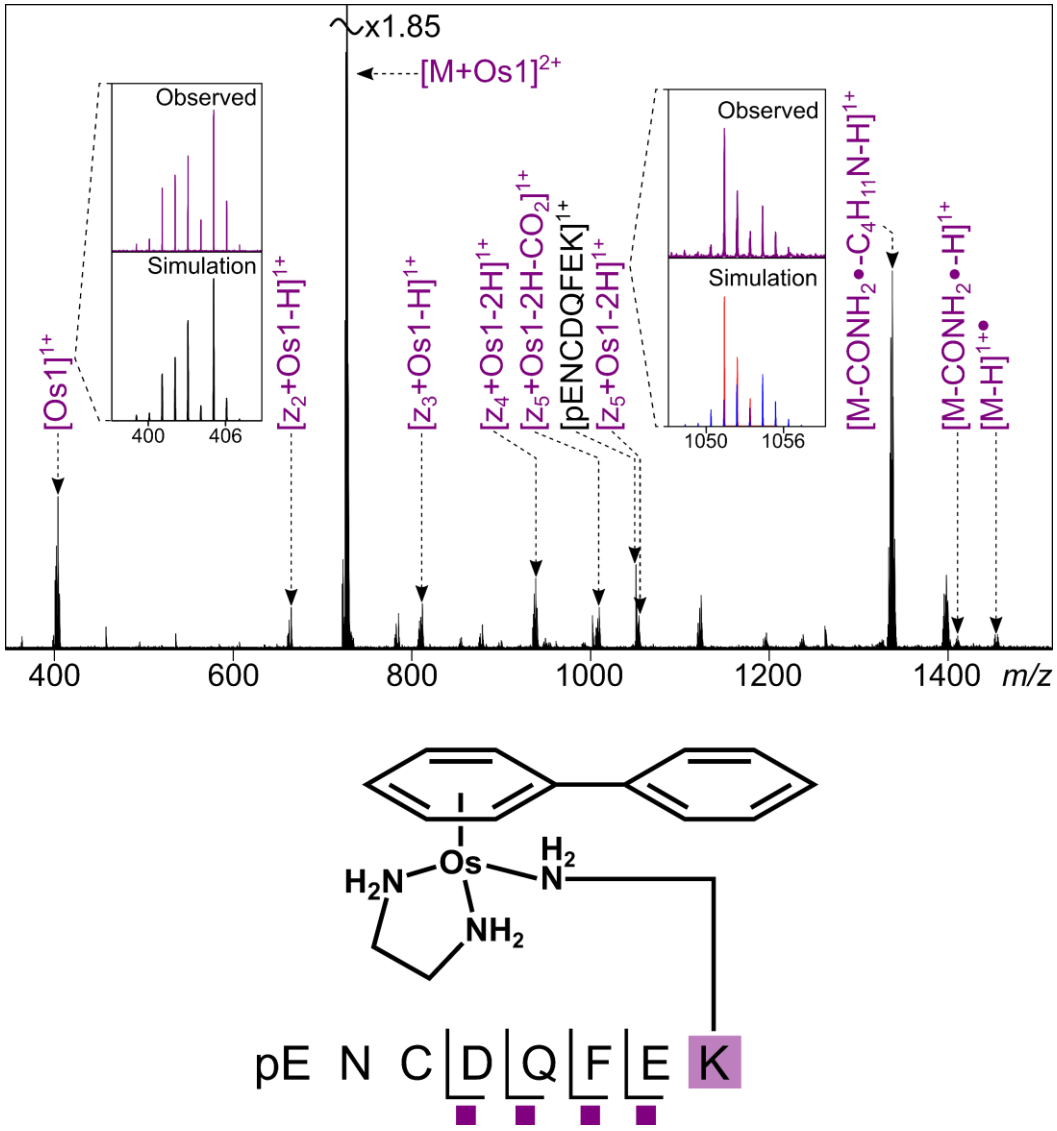


Figure 2.7 ECD MS/MS of the osmiated species, 728.25912 *m/z*, resulting from the reaction of Os1 with pENCDQFEK (residue 413-420). Osmium containing species are labelled in purple. Inset on the left shows the dissociation of the osmium complex, Os1, and inset on the right shows an overlapping isotope pattern by the unmodified peptide $[pENCDQFEK]^{1+}$ (shown in red) and $[z_5+Os1-2H]^{1+}$ (shown in blue).

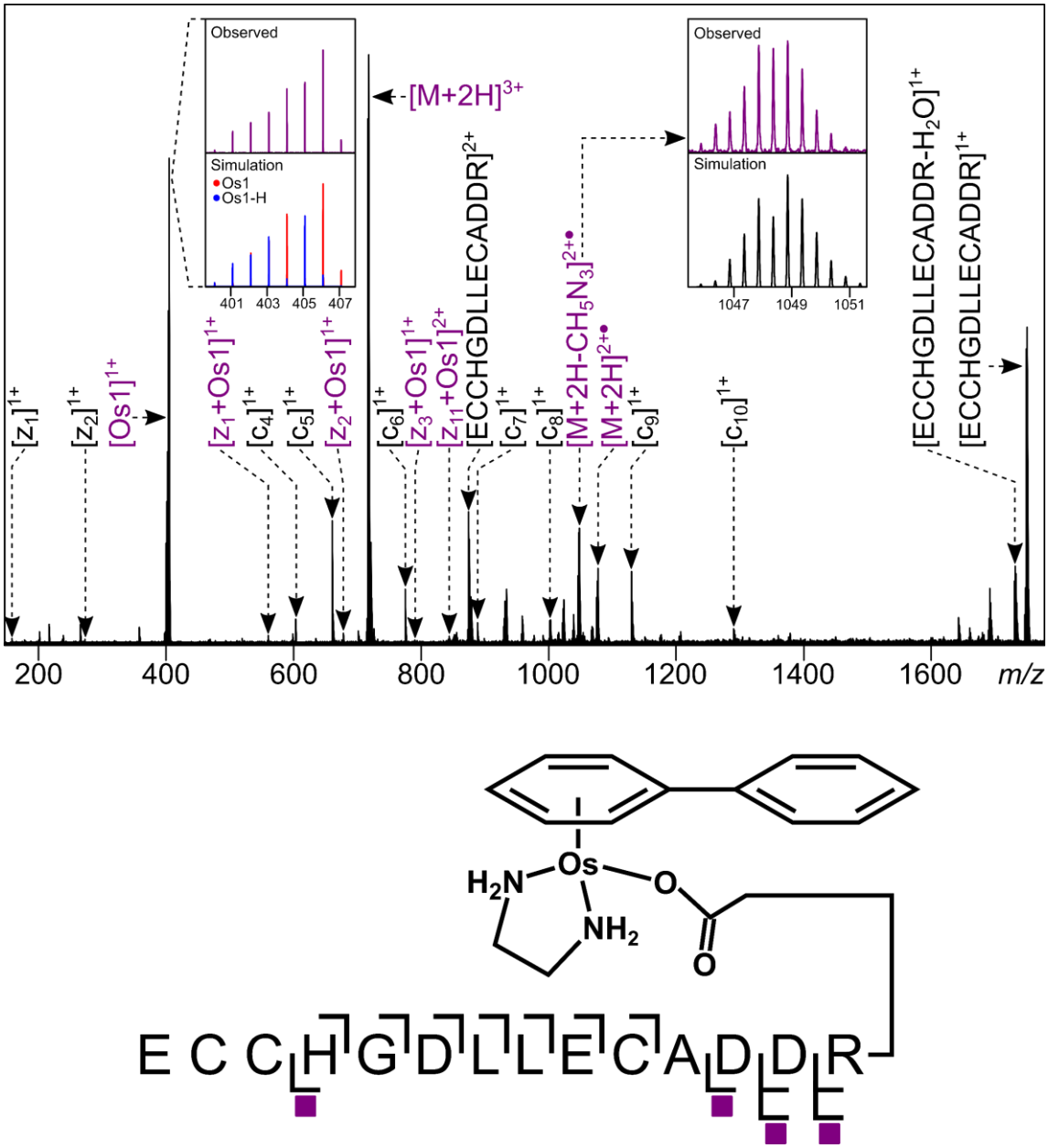


Figure 2.8 ECD MS/MS of the osmiated species, 718.58996 *m/z*. Osmium containing species are labelled in purple. Inset on the left shows an overlap distribution of the osmium complex, Os1 is shown in red and Os1-H is shown in blue. Inset on the right shows a unique side chain loss of the arginine residue from the charge-reduced species.

The peak with 718.58996 *m/z* is another osmiated species (ECCHGDLLECADDR+Os1) picked out by the SNAP-LC algorithm and the ECD

MS/MS spectrum is shown below in Figure 2.8 (Table S2.5). From the inset of the left, it shows an overlap isotope distribution of two osmium complexes, Os1 ($C_{14}H_{18}N_2Os$, shown with red simulation) and Os1-H ($C_{14}H_{17}N_2Os$, shown with blue simulation). The unmodified peptide (ECCHGDLLECADDR) (residue 267-280) was observed. However, it is important to investigate whether the peptide binds to Os1 or Os1-H. The osmium containing fragments have narrowed the binding site to the arginine residue (R) and the C-terminus. There is one important species in the ECD MS/MS spectrum that allows the determination of the exact binding site, whether the osmium complex is binding to the arginine side chain or the C-terminus. The unique loss of arginine side chain (CH_5N_3, -59.048 Da) was observed from the charge-reduced species (CRS).[40] It is worth noting that the species still possesses a distinct osmium isotope pattern, suggesting that the loss of the arginine side chain is not influencing the observation of osmium binding. Therefore, the peptide is binding to the osmium complex *via* the C-terminal carboxylate group (COO^-), leaving the osmium complex at a charge of 1+, and the remaining two positive charges are contributed by 2 protons on the peptide.

Figure 2.9 shows another osmiated peptide at 663.79464 *m/z*, and two osmium complexes with different composition, Os1 and Os1-H were observed in the spectrum again. The fragmentation pattern localises the binding site to glutamic acid (E) and valine (V) in the peptide sequence (Table S2.6). However, valine is not a viable binding location because of the purely hydrocarbon side chain, therefore, the peptide is binding to the osmium complex *via* the carboxylate side chain of glutamic acid. The interaction leaves the osmium complex at a 1+ charge and together with a proton from the peptide, allowed the observation of the osmiated peptide precursor as a doubly charged species.

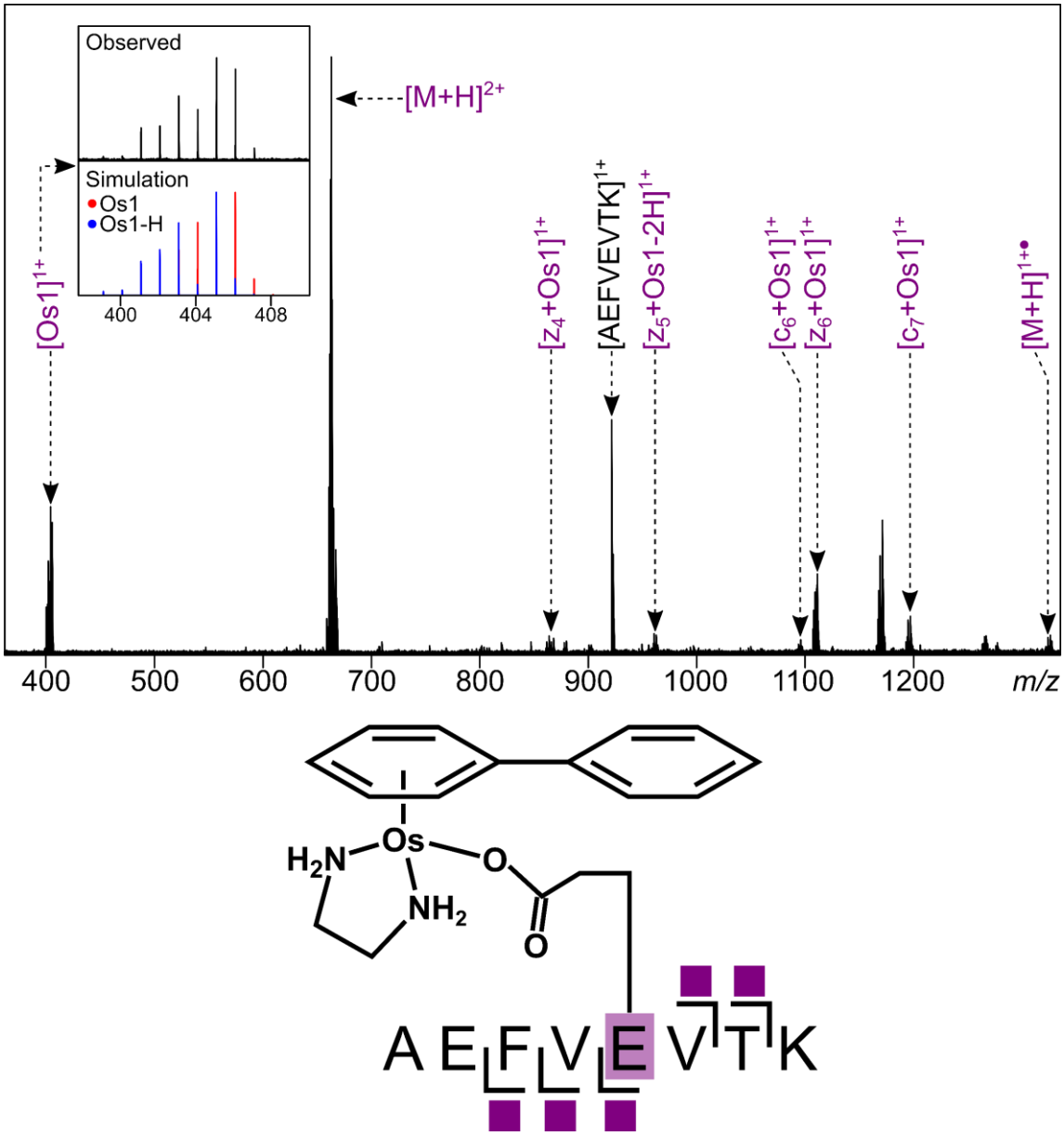


Figure 2.9 ECD MS/MS of the osmiated species, 663.79464 *m/z*. Osmium containing species are labelled in purple. Inset shows an overlap distribution of the osmium complex, Os1 is shown in red and Os1-H is shown in blue.

From the peak list generated by the SNAP-LC algorithm, a series of nLC-ECD MS/MS experiments were performed and are summarised Figure 2.10. Even though a separation process was performed using a RP C18 column prior to MS analysis, ion suppression and co-elution were still potential obstacles to obtain effective isolation and fragmentation. And unfortunately, this is exactly what was observed with the nLC-

MS/MS experiments. To overcome this problem, a truly data-independent acquisition method, such as two dimensional mass spectrometry,[41] would be required, so that all observed precursor can be fragmented and characterised.

Osmiated Peptides

m/z	Charge State	MS/MS ?	Sequence/Remarks
427.8100	3+	No	below threshold
430.8431	3+	Yes	131 - DDSPDLPK - 138
467.6768	4+	No	76 - TCVADESHAGCEK - 88, below threshold
547.2157	2+	Yes	236 - AWSVAR - 241
616.5847	3+	Yes	286 - YICDNQDTISSK - 297
623.2360	3+	Yes	76 - TCVADESHAGCEK - 88
628.2153	3+	Yes	106 - ETYGDMADCCEK - 117
633.5485	3+	No	below threshold
636.2423	3+	Yes	375 - EYEATLEECCAK - 386
641.2140	2+	Yes	-
645.7638	2+	No	131 - DDSPDLPK - 138
663.7950	2+	Yes	249 - AEFVEVTK - 256
718.5887	3+	Yes	267 - ECCHGDLLECADDR - 280
728.2591	2+	Yes	413 - pENCDQFEK - 420
763.2837	2+	Yes	-
777.2912	2+	Yes	-
798.2844	2+	No	-
867.2759	2+	Yes	156 - KFWGK - 160
941.8183	2+	No	106 - ETYGDMADCCEK - 117
953.8581	2+	No	375 - EYEATLEECCAK - 386

Figure 2.10 Osmiated BSA peptides identified from the SNAP-LC analysis of the nLC-MS data. The identified binding locations are highlighted in purple.

Conclusions

Microwave digestion was found to have digested BSA extensively into short peptides, yielding dominantly singly-charged species compared to standard tryptic digestion, which mainly multiply-charged species were observed. Parameters of microwave digestion can be further optimised, *e.g.* lowering down the power and shorter hold time.

Os1-Cl was reacted with a tryptic digested BSA as a model system to develop a methodology for metallodrug-binding site MS analysis. Using ultra-high resolution mass spectrometry, the metallodrug Os1-Cl was shown to react with the tryptic digested BSA protein and produced a series of osmiated species. With direct infusion, a few

osmiated peptides were observed in low intensities possibly due to ion suppression and space charge effects from the high abundance unmodified peptides.

With CAD MS/MS, the osmium complex dissociated readily even using low collision energy (5 V), leaving with the unmodified peptide, suggesting the binding interaction between the osmium complex and the target peptide is weaker than the protein backbone. ECD MS/MS was shown to preserve the osmium complex on the target peptides upon electron-based fragmentation, and more binding information was revealed. Os1-Cl was found to interact with the side chain of lysine of peptide KFWGK (156-160) as well as a weak π-π interaction with the aromatic region of tryptophan.

nLC-MS and nLC-MS/MS were also attempted to separate the complex mixture of tryptic digested peptides (unmodified and osmiated), so that more osmiated species can be observed and studied. SNAP-LC algorithm was used to pick out any species containing an osmium modification,[35] and generated a prefer list for targeted nLC-ECD MS/MS experiments, to study the exact binding location of the osmium complex in the target peptides. More osmiated peptides with higher ion intensities were indeed observed and allowed more effective isolation and fragmentation compared to direct infusion, binding sites were revealed to be the side chains of glutamic acids and lysines, as well as the N- and C- termini.

This work demonstrates the flexibility of FT-ICR mass spectrometry and its capability to employ different fragmentation techniques to characterise metallodrug binding to biomolecules of interest. FT-ICR MS can also be coupled to hyphenated techniques, such as liquid chromatography (LC) or gas chromatography (GC) for analyte separation prior to MS analysis. FT-ICR MS provides the highest resolving power and mass accuracy performance of any MS platform available, offer extremely high confident spectral assignment, with mass errors often measured to sub-ppm or even ppb. Therefore, FT-ICR MS is particularly adept for the characterisation of metallodrug-biomolecule interactions.

References

(1) Zhang, P.; Sadler, P. J. Advances in the Design of Organometallic Anticancer Complexes. *J. Organomet. Chem.* **2017**, *839*, 5–14.

(2) Wootton, C. A.; Sanchez-Cano, C.; Liu, H.-K.; Barrow, M. P.; Sadler, P. J.; O'Connor, P. B. Binding of an Organo-Osmium(II) Anticancer Complex to Guanine and Cytosine on DNA Revealed by Electron-Based Dissociations in High Resolution Top-Down FT-ICR Mass Spectrometry. *Dalton Trans.* **2015**, *44* (8), 3624–3632.

(3) Dyson, P. J.; Sava, G. Metal-Based Antitumour Drugs in the Post Genomic Era. *Dalt. Trans.* **2006**, *35* (16), 1929–1933.

(4) Kelland, L. The Resurgence of Platinum-Based Cancer Chemotherapy. *Nat. Rev. Cancer* **2007**, *7*, 573–584.

(5) Gras, M.; Therrien, B.; Süss-Fink, G.; Casini, A.; Edafe, F.; Dyson, P. J. Anticancer Activity of New Organo-Ruthenium, Rhodium and Iridium Complexes Containing the 2-(Pyridine-2-Yl)Thiazole N,N-Chelating Ligand. *J. Organomet. Chem.* **2010**, *695* (8), 1119–1125.

(6) Wills, R. H.; Habtemariam, A.; Lopez-Clavijo, A. F.; Barrow, M. P.; Sadler, P. J.; O'Connor, P. B. Insights into the Binding Sites of Organometallic Ruthenium Anticancer Compounds on Peptides Using Ultra-High Resolution Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2014**, *25* (4), 662–672.

(7) Tian, Z.; Li, J.; Zhang, S.; Xu, Z.; Yang, Y.; Kong, D.; Zhang, H.; Ge, X.; Zhang, J.; Liu, Z. Lysosome-Targeted Chemotherapeutics: Half-Sandwich Ruthenium(II) Complexes That Are Selectively Toxic to Cancer Cells. *Inorg. Chem.* **2018**, *57*, 10498–10502.

(8) Huang, H.; Yu, B.; Zhang, P.; Huang, J.; Chen, Y.; Gasser, G.; Ji, L.; Chao, H. Highly Charged Ruthenium(II) Polypyridyl Complexes as Lysosome-Localized Photosensitizers for Two-Photon Photodynamic Therapy. *Angew. Chemie - Int. Ed.* **2015**, *54* (47), 14049–14052.

(9) Nam, J. S.; Kang, M. G.; Kang, J.; Park, S. Y.; Lee, S. J. C.; Kim, H. T.; Seo, J. K.; Kwon, O. H.; Lim, M. H.; Rhee, H. W.; et al. Endoplasmic Reticulum-Localized Iridium(III) Complexes as Efficient Photodynamic Therapy Agents *via* Protein Modifications. *J. Am. Chem. Soc.* **2016**, *138* (34), 10968–10977.

(10) Wootton, C. A.; Millett, A. J.; Lopez-Clavijo, A. F.; Chiu, C. K. C.; Barrow, M. P.; Clarkson, G. J.; Sadler, P. J.; O'Connor, P. B. Structural Analysis of Peptides Modified with Organo-Iridium Complexes, Opportunities from Multi-Mode Fragmentation. *Analyst* **2019**, *144*, 1575–1581.

(11) Huang, H.; Zhang, P.; Qiu, K.; Huang, J.; Chen, Y.; Ji, L.; Chao, H. Mitochondrial Dynamics Tracking with Two-Photon Phosphorescent Terpyridyl Iridium(III) Complexes. *Sci. Rep.* **2016**, *6* (Iii), 20887.

(12) Zhang, P.; Chiu, C. K. C.; Huang, H.; Lam, Y. P. Y.; Habtemariam, A.; Malcomson, T.; Paterson, M. J.; Clarkson, G. J.; O'Connor, P. B.; Chao, H.; et al. Organoiridium Photosensitizers Induce Specific Oxidative Attack on Proteins within Cancer Cells. *Angew. Chemie Int. Ed.* **2017**, 14898–14902.

(13) Needham, R. J.; Sanchez-Cano, C.; Zhang, X.; Romero-Canelón, I.; Habtemariam, A.; Cooper, M. S.; Meszaros, L.; Clarkson, G. J.; Blower, P. J.; Sadler, P. J. In-Cell Activation of Organo-Osmium(II) Anticancer Complexes. *Angew. Chemie - Int. Ed.* **2017**, *56* (4), 1017–1020.

(14) Hearn, J. M.; Romero-Canelón, I.; Munro, A. F.; Fu, Y.; Pizarro, A. M.; Garnett, M. J.; McDermott, U.; Carragher, N. O.; Sadler, P. J. Potent Organo-Osmium Compound Shifts Metabolism in Epithelial Ovarian Cancer Cells. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112* (29), E3800–E3805.

(15) Sanchez-Cano, C.; Romero-Canelón, I.; Yang, Y.; Hands-Portman, I. J.; Bohic, S.; Cloetens, P.; Sadler, P. J. Synchrotron X-Ray Fluorescence Nanoprobe Reveals Target Sites for Organo-Osmium Complex in Human Ovarian Cancer Cells. *Chem. - A Eur. J.* **2017**, *23* (11), 2512–2516.

(16) Banerjee, S.; Soldevila-Barreda, J. J.; Wolny, J. A.; Wootton, C. A.;

Habtemariam, A.; Romero-Canelon, I. New Activation Mechanism for Half-Sandwich Organometallic Anticancer Complexes. *Chem. Sci.* **2018**, *9*, 3177–3185.

(17) Serna, A.; Galán-Cobo, A.; Rodrigues, C.; Sánchez-Gomar, I.; Toledo-Aral, J. J.; Moura, T. F.; Casini, A.; Soveral, G.; Echevarría, M. Functional Inhibition of Aquaporin-3 With a Gold-Based Compound Induces Blockage of Cell Proliferation. *J. Cell. Physiol.* **2014**, *229* (11), 1787–1801.

(18) Zou, T.; Lum, C. T.; Lok, C. N.; To, W. P.; Low, K. H.; Che, C. M. A Binuclear Gold(I) Complex with Mixed Bridging Diphosphine and Bis(N-Heterocyclic Carbene) Ligands Shows Favorable Thiol Reactivity and Inhibits Tumor Growth and Angiogenesis in Vivo. *Angew. Chemie - Int. Ed.* **2014**, *53* (23), 5810–5814.

(19) Venkatesh, V.; Wedge, C. J.; Romero-Canelón, I.; Habtemariam, A.; Sadler, P. J. Spin-Labelled Photo-Cytotoxic Diazido Platinum(IV) Anticancer Complex. *Dalt. Trans.* **2016**, *45* (33), 13034–13037.

(20) Li, H.; Lin, T. Y.; Van Orden, S. L.; Zhao, Y.; Barrow, M. P.; Pizarro, A. M.; Qi, Y.; Sadler, P. J.; O'Connor, P. B. Use of Top-down and Bottom-up Fourier Transform Ion Cyclotron Resonance Mass Spectrometry for Mapping Calmodulin Sites Modified by Platinum Anticancer Drugs. *Anal. Chem.* **2011**, *83* (24), 9507–9515.

(21) Mügge, C.; Rothenburger, C.; Beyer, A.; Görls, H.; Gabbiani, C.; Casini, A.; Michelucci, E.; Landini, I.; Nobili, S.; Mini, E.; et al. Structure, Solution Chemistry, Antiproliferative Actions and Protein Binding Properties of Non-Conventional Platinum(II) Compounds with Sulfur and Phosphorus Donors. *Dalton Trans.* **2011**, *40* (9), 2006–2016.

(22) Wootton, C. A.; Sanchez-Cano, C.; Lopez-Clavijo, A. F.; Shaili, E.; Barrow, M. P.; Sadler, P. J.; O'Connor, P. B. Sequence-Dependent Attack on Peptides by Photoactivated Platinum Anticancer Complexes. *Chem. Sci.* **2018**, *9* (10), 2733–2739.

(23) Jin, C.; Guan, R.; Wu, J.; Yuan, B.; Wang, L.; Huang, J.; Wang, H.; Ji, L.; Chao, H. Rational Design of NIR-Emitting Iridium(III) Complexes for Multimodal Phosphorescence Imaging of Mitochondria under Two-Photon Excitation. *Chem. Commun.* **2017**, *53*, 10374–10377.

(24) Hearn, J. M.; Romero-Canelón, I.; Qamar, B.; Liu, Z.; Hands-Portman, I.; Sadler, P. J. Organometallic Iridium(III) Anticancer Complexes with New Mechanisms of Action: NCI-60 Screening, Mitochondrial Targeting, and Apoptosis. *ACS Chem. Biol.* **2013**, *8* (6), 1335–1343.

(25) Vinje, J.; Sletten, E. NMR Spectroscopy of Anticancer Platinum Drugs. *Anticancer. Agents Med. Chem.* **2007**, *7*, 35–54.

(26) Handing, K. B.; Niedzialkowska, E.; Shabalin, I. G.; Kuhn, M. L.; Zheng, H.; Minor, W. Characterizing Metal-Binding Sites in Proteins with X-Ray Crystallography. *Nat. Protoc.* **2018**, *13* (5), 1062–1090.

(27) Jennings, K. R. Colllision-Induced Decompositions of Aromatic Ions. *Int. J. Mass Spectrom. Ion Phys.* **1968**, *1* (3), 227–235.

(28) Zubarev, R.; Kelleher, N. L.; McLafferty, F. W. Electron Capture Dissociation of Multiply Charged Protein Cations. A Nonergodic Process. *J. Am. Chem. Soc* **1998**, *120* (16), 3265–3266.

(29) Little, D. P.; Speir, J. P.; Senko, M. W.; O'Connor, P. B.; McLafferty, F. W. Infrared Multiphoton Dissociation of Large Multiply Charged Ions for Biomolecule Sequencing. *Anal. Chem.* **1994**, *66* (18), 2809–2815.

(30) Marshall, A. G.; Hendrickson, C. L.; Jackson, G. S. Fourier Transform Ion Cyclotron Resonance Mass Spectrometry: A Primer. *Mass Spectrom. Rev.* **1998**, *17* (1), 1–35.

(31) Amster, I. J. Fourier Transform Mass Spectrometry. *J Mass Spectrom* **1996**, *31* (September), 1325–1337.

(32) Li, H.; Wolff, J. J.; Van Orden, S. L.; Loo, J. A. Native Top-down Electrospray Ionization-Mass Spectrometry of 158 KDa Protein Complex by High-Resolution

Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Anal. Chem.* **2014**, *86* (1), 317–320.

(33) Peacock, A. F. A.; Habtemariam, A.; Fernández, R.; Walland, V.; Fabbiani, F. P. A.; Parsons, S.; Aird, R. E.; Jodrell, D. I.; Sadler, P. J. Tuning the Reactivity of Osmium(II) and Ruthenium(II) Arene Complexes under Physiological Conditions. *J. Am. Chem. Soc.* **2006**, *128* (5), 1739–1748.

(34) Caravatti, P.; Allemann, M. The 'Infinity Cell': A New Trapped-ion Cell with Radiofrequency Covered Trapping Electrodes for Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Org. Mass Spectrom.* **1991**, *26* (5), 514–518.

(35) Wootton, C. A.; Lam, Y. P. Y.; Willetts, M.; van Agthoven, M. A.; Barrow, M. P.; Sadler, P. J.; O′Connor, P. B. Automatic Assignment of Metal-Containing Peptides in Proteomic LC-MS and MS/MS Data Sets. *Analyst* **2017**, *142* (11), 2029–2037.

(36) Zhang, L.; Rempel, D.; Pramanik, B. N.; Gross, M. L. Accurate Mass Measurements by Fourier Transform Mass Spectrometry. *Mass Spectrom. Rev.* **2005**, *24*, 286–309.

(37) Majorek, K. A.; Porebski, P. J.; Dayal, A.; Zimmerman, M. D.; Jablonska, K.; Stewart, A. J.; Chruszcz, M.; Minor, W. Structural and Immunologic Characterization of Bovine, Horse, and Rabbit Serum Albumins. *Mol. Immunol.* **2012**, *52* (3–4), 174–182.

(38) Ledford, E. B.; Rempel, D. L.; Gross, M. L. Space Charge Effects in Fourier Transform Mass Spectrometry. Mass Calibration. *Anal. Chem.* **1984**, *56* (14), 2744–2748.

(39) Fauchere, J. L.; Pliska, V. Hydrophobic Parameters Pi of Amino Acid Side Chains from the Partitioning of N-Acetyl-Amino Acid Amides. *Eur. J. Med. Chem.* **1983**, *18* (4), 369–375.

(40) Cooper, H. J.; Hudgins, R. R.; Håkansson, K.; Marshall, A. G. Characterization

of Amino Acid Side Chain Losses in Electron Capture Dissociation. *J. Am. Soc. Mass Spectrom.* **2002**, *13* (3), 241–249.

(41) van Agthoven, M. A.; Lam, Y. P. Y.; O'Connor, P. B.; Rolando, C.; Delsuc, M. A. Two-Dimensional Mass Spectrometry: New Perspectives for Tandem Mass Spectrometry. *Eur. Biophys. J.* **2019**, *48* (3), 213–229.

Supplementary Information

Table S2.1 Assignments for Figure 2.2: Full MS of [Os1-Cl]-BSA. Osmiated species are highlighted in purple.

Assignment	Charge State	Elemental Composition	Theoretical Mass (m/z)	Observed Mass (m/z)	Mass Error (ppm)
[P249-256]	2+	$C_{42}H_{69}N_9O_{14}$	461.747638	461.747652	0.03
[P37-44]	2+	$C_{43}H_{65}N_{11}O_{15}$	487.732538	487.732863	0.67
[P337-340]	1+	$C_{20}H_{37}N_6O_8S$	521.238800	521.238954	0.30
[P157-160]	1+	$C_{28}H_{37}N_6O_5$	537.282000	537.281730	-0.50
[P437-451]	3+	$C_{72}H_{129}N_{20}O_{23}$	547.317417	547.316932	-0.89
[P499-507]	2+	$C_{43}H_{77}N_{15}O_{17}S_2$	569.752638	569.752840	0.35
[P257-266]	2+	$C_{52}H_{94}N_{14}O_{15}$	577.350588	577.350358	-0.40
[P66-75]	2+	$C_{53}H_{88}N_{12}O_{17}$	582.318988	582.318997	0.02
[P25-34]	2+	$C_{49}H_{82}N_{18}O_{17}$	597.304738	597.304625	-0.19
[P223-232]	2+	$C_{50}H_{84}N_{16}O_{16}S$	598.298038	598.298125	0.15
[P35-44]	2+	$C_{58}H_{86}N_{14}O_{17}$	625.314238	625.314125	-0.18
[P236-241]	1+	$C_{31}H_{49}N_{10}O_8$	689.372900	689.372881	-0.03
[P223-228]	1+	$C_{28}H_{52}N_9O_{10}S$	706.355200	706.355128	-0.10
[P106-122]	3+	$C_{83}H_{131}N_{24}O_{35}S_3$	706.611817	706.612325	0.72
[P400-412]	2+	$C_{70}H_{121}N_{19}O_{20}$	773.951238	773.951128	-0.14
[P257-263]	1+	$C_{35}H_{65}N_8O_{12}$	789.471600	789.471689	0.11
[P437-451]	2+	$C_{72}H_{128}N_{20}O_{23}$	820.472488	820.472406	-0.10
[P156-160]	2+	$C_{82}H_{114}N_{18}O_{12}Os$	867.423852	867.423896	0.05
[P249-256]	1+	$C_{42}H_{68}N_9O_{14}$	922.488000	922.488128	0.14
[P161-167]	1+	$C_{44}H_{67}N_{10}O_{12}$	927.493400	927.493654	0.27
[P341-346]	2+	$C_{78}H_{115}N_{20}O_{24}Os$	953.900301	953.900952	0.68
[P598-607]	1+	$C_{44}H_{80}N_{11}O_{15}$	1002.583000	1002.582965	-0.03
[P549-557]	1+	$C_{46}H_{84}N_{11}O_{14}$	1014.619400	1014.618565	-0.82
[P310-318]	1+	$C_{44}H_{74}N_{13}O_{16}S$	1072.509200	1072.508912	-0.27
				Average	-0.01 ppm
				Absolute Average	0.37 ppm
				Standard Deviation	0.41 ppm

Table S2.2 Assignments for Figure 2.3: CAD MS/MS of the osmiated species at 867.42385 *m/z*. Osmiated species are highlighted in purple.

Assignment	Charge State	Elemental Composition	Theoretical Mass (m/z)	Observed Mass (m/z)	Mass Error (ppm)
[Os1]	1+	$C_{14}H_{18}N_2Os$	406.107956	406.10795	-0.02
[P156-160]	2+	$C_{82}H_{114}N_{18}O_{12}Os$	867.423852	867.42377	-0.10
[P156-160]	2+	$C_{68}H_{97}N_{16}O_{12}$	1329.746639	1329.74684	0.15
				Average	0.01 ppm
				Absolute Average	0.02 ppm
				Standard Deviation	0.09 ppm

Table S2.3 Assignments for Figure 2.4: ECD MS/MS of the osmiated species at 867.42385 *m/z*. Osmiated species are highlighted in purple.

Assignment	Charge State	Elemental Composition	Theoretical Mass (m/z)	Observed Mass (m/z)	Mass Error (ppm)
[Os1]	1+	$C_{14}H_{17}N_2Os$	405.100131	405.10012	-0.03
[FWGK]	1+	$C_{28}H_{37}N_6O_5$	537.281995	537.28188	-0.22
[(GK)+Os1]	1+	$C_{22}H_{18}N_5O_3Os$	592.101967	592.10207	0.17
$[(KFWGK)_2$+Os1]	2+	$C_{82}H_{114}N_{18}O_{12}Os$	867.423852	867.42406	0.24
[(GK+KFWGK)+Os1]	1+	$C_{56}H_{83}N_{13}O_9Os$	1273.605078	1273.60603	0.74
[(WGK+KFWGK)+Os1]	1+	$C_{67}H_{85}N_{15}O_{10}Os$	1451.621978	1451.62111	-0.60
[(KFWG+KFWGK)+Os1]	1+	$C_{76}H_{95}N_{15}O_{11}Os$	1585.695326	1585.69412	-0.76
				Average	-0.07 ppm
				Absolute Average	0.39 ppm
				Standard Deviation	0.47 ppm

Table S2.4 Assignments for Figure 2.7: ECD MS/MS of the osmiated species at 728.25912 *m/z*. Osmiated species are highlighted in purple.

Assignment	Charge State	Elemental Composition	Theoretical Mass (m/z)	Observed Mass (m/z)	Mass Error (ppm)
Os1	1+	$C_{14}H_{17}N_2Os_1$	405.100131	405.10013	0.00
z2+Os1-H	1+	$C_{25}H_{37}N_4O_5Os_1$	665.237429	665.23686	-0.86
M	2+	$C_{57}H_{80}N_{14}O_{17}Os_1S_1$	728.257656	728.25759	-0.09
z3+Os1-H	1+	$C_{34}H_{46}N_5O_6Os_1$	812.305927	812.30578	-0.18
z4+Os1-2H	1+	$C_{39}H_{53}N_7O_8Os_1$	939.356737	939.35667	-0.07
z5+Os1-2H-CO_2	1+	$C_{42}H_{58}N_8O_9Os_1$	1010.393890	1010.39390	0.01
pENCDQFEK	1+	$C_{43}H_{63}N_{12}O_{17}S_1$	1051.414936	1051.41476	-0.17
z5+Os1-2H	1+	$C_{43}H_{58}N_8O_{11}Os_1$	1054.383739	1054.38340	-0.32
M-$CONH_2$•-$C_4H_{11}N$-H	1+	$C_{52}H_{66}N_{12}O_{16}Os_1S_1$	1338.405171	1338.40618	0.75
M-$CONH_2$•-H	1+	$C_{56}H_{77}N_{13}O_{16}Os_1S_1$	1411.494381	1411.49488	0.35
M-H	1+•	$C_{57}H_{79}N_{14}O_{17}Os_1S_1$	1455.508034	1455.50860	0.39
				Average	-0.02 ppm
				Absolute Average	0.29 ppm
				Standard Deviation	0.42 ppm

Table S2.5 Assignments for Figure 2.8: ECD MS/MS of the osmiated species at 718.58996 *m/z*. Osmiated species are highlighted in purple.

Assignment	Charge State	Elemental Composition	Theoretcal Mass (m/z)	Observed Mass (m/z)	Mass Error (ppm)
z1	1+	$C_6H_{13}N_3O_2$	159.100228	159.10023	0.01
z2	1+	$C_{10}H_{18}N_4O_5$	274.127172	274.12717	-0.01
z1+Os1	1+	$C_{20}H_{27}N_5O_2Os_1$	561.177478	561.17716	-0.57
c4	1+	$C_{21}H_{34}N_9O_8S_2$	604.196629	604.19663	0.00
c5	1+	$C_{23}H_{37}N_{10}O_9S_2$	661.218093	661.21759	-0.76
z2+Os1	1+	$C_{24}H_{35}N_6O_5Os_1$	679.227915	679.22725	-0.98
[M+2H]	3+	$C_{80}H_{123}N_{24}O_{28}Os_1S_3$	718.590039	718.58996	-0.11
c6	1+	$C_{27}H_{42}N_{11}O_{12}S_2$	776.245037	776.24456	-0.61
z3+Os1	1+	$C_{28}H_{37}N_7O_8Os_1$	791.231421	791.23162	0.25
z11+Os1	2+	$C_{65}H_{97}N_{18}O_{21}S_1Os_1$	844.820213	844.82040	0.22
ECCHGDLLECADDR	2+	$C_{66}H_{106}N_{22}O_{28}S_3$	875.334903	875.33445	-0.52
c7	1+	$C_{33}H_{53}N_{12}O_{13}S_2$	889.329101	889.32869	-0.46
c8	1+	$C_{39}H_{64}N_{13}O_{14}S_2$	1002.413165	1002.41276	-0.40
CRS-CH_5N_3	2+	$C_{79}H_{119}N_{21}O_{28}Os_1S_3$	1048.865073	1048.86508	0.01
c9	1+	$C_{44}H_{71}N_{14}O_{17}S_2$	1131.455759	1131.45678	0.90
c10	1+	$C_{49}H_{79}N_{16}O_{19}S_3$	1291.486408	1291.48595	-0.35
ECCHGDLLECADDR-H_2O	1+	$C_{66}H_{103}N_{22}O_{27}S_3$	1731.651966	1731.65112	-0.49
ECCHGDLLECADDR	1+	$C_{66}H_{105}N_{22}O_{28}S_3$	1749.662530	1749.66201	-0.30
				Average	-0.23 ppm
				Absolute Average	0.39 ppm
				Standard Deviation	0.44 ppm

Table S2.6 Assignments for Figure 2.9: ECD MS/MS of the osmiated species at 663.79464 *m/z*. Osmiated species are highlighted in purple.

Assignment	Charge State	Elemental Composition	Theoretical Mass (m/z)	Observed Mass (m/z)	Mass Error (ppm)
Os1	1+	$C_{14}H_{17}N_2Os_1$	405.100131	405.09994	-0.47
M+H	2+	$C_{56}H_{85}N_{11}O_{14}Os_1$	663.794317	663.79464	0.49
z4+Os1	1+	$C_{34}H_{54}N_6O_8Os_1$	866.361435	866.36071	-0.84
AEFVEVTK	1+	$C_{42}H_{68}N_9O_{14}$	922.488024	922.48859	0.61
z5+Os1-2H	1+	$C_{39}H_{61}N_7O_9Os_1$	963.414258	963.41465	0.41
c6+Os1	1+	$C_{46}H_{66}N_9O_{10}Os_1$	1096.454534	1096.45506	0.48
z6+Os1	1+	$C_{48}H_{71}N_8O_{10}Os_1$	1111.490620	1111.49097	0.31
c7+Os1	1+	$C_{50}H_{73}N_{10}O_{12}Os_1$	1197.502276	1197.50269	0.35
M+H	1+•	$C_{56}H_{85}N_{11}O_{14}Os_1$	1326.581356	1326.58186	0.38
				Average	0.19 ppm
				Absolute Average	0.48 ppm
				Standard Deviation	0.50 ppm

Chapter 3

Reduction of Disulfide Bonds in Proteins for Enhanced MS and MS/MS Capabilities

Disulfide bonds have important biological functions but also present a distinct challenge in tandem mass spectrometry analysis of proteins. Chemical reduction and alkylation are often used to overcome this hurdle in bottom-up or top-down approaches, however, they often induce undesired modifications or present limited effectiveness by the sample conditions, such as pH environment which is non-optimum for reducing disulfide bonds.

This chapter of the thesis is centred around the study of alternative reduction techniques, without the need of introducing extra chemicals into the sample for reduction. These alternative methods have shown to effectively reduce the disulfide bonds, yielding better fragmentation efficiencies, and allow better protein identification and characterisation with tandem mass spectrometry, allowing future studies of metallodrug modified proteins without chemical digestion.

Although the results presented within this chapter are predominantly based on FT-ICR MS and all results were analysed and written by the thesis Author, this was a collaborative project with researchers, namely; Christopher Wootton assisted with some of the data acquisition and helped with analysing the complex data. Antec Scientific kindly provided the ROXY electrochemical system for this electrochemical reduction study, which was later published as an application note for Antec, titled “Monitoring the Electrochemical Reduction of Disulfide Bonds in Proteins for Enhanced MS and MS/MS Capabilities” and can be found on the Antec Scientific webpage: “https://antecscientific.com/support/documents-and-downloads/application-notes” by Christopher A. Wootton and Cookson Chiu, Agnieszka Kraj, Thomas Lambert, Simon Lambert, and Peter B. O’Connor.

A manuscript entitled “Reduction of Disulfide Bonds in Proteins for Enhanced MS and MS/MS Capabilities” by Cookson K. C. Chiu and Christopher A. Wootton, Mark P. Barrow, and Peter B. O’Connor, has been prepared for submission to the

Journal of the American Society for Mass Spectrometry (JASMS) based directly on the contents and findings of this chapter.

Abstract

Disulfide bonds are an important class of protein post-translational modifications and involved in many vital biological functions, *e.g.* protein folding and tertiary structure. However, the presence of disulfide bonds in proteins often present as a challenge in MS/MS-based experiments. Fragmentation efficiency in tandem mass spectrometry (MS/MS) is often hindered by the presence of "disulfide bond protected" regions, by creating cyclic structures along the protein backbone. Chemical reduction and alkylation are often used to reduce these disulfide linkages and "open up" the protein structure, allowing more protonation, thus, achieving higher charge states. However, by inducing chemical-based reducing reagents into the sample can easily disrupt many chemically labile modifications or worse, induce modifications onto the proteins artificially, increasing the difficulty of data analysis greatly. Alternative reduction methods, such as electrochemical reduction and radical-induced reduction, have been investigated to cleave these cystine linkages without traditional techniques, preventing unnecessary chemical disruptions of the sample. Herein, the performance and capability of electrochemical reduction and radical-induced reduction were investigated by using medium-sized proteins, including lysozyme, ribonuclease A, and insulin. The reduction processes of both methods were monitored using ultra-high resolution mass spectrometry, with the aid of various fragmentation techniques. The results show that both methods can effectively reduce the disulfide bonds and enhance the fragmentation efficiency, allowing better protein identification and characterisation, and present possible methods for studying proteins modified with chemically labile species, such as chemotherapeutic metallodrugs.

Introduction

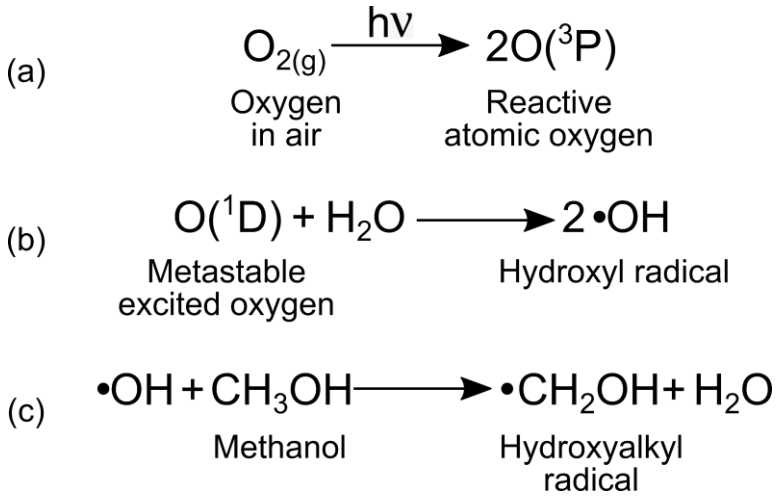
Proteomics is a popular research topic worldwide and many branches of research are very active, including protein identification,[1] protein quantification,[2,3] post-translation modification (PTM) characterisation,[4,5] and even protein structure elucidation.[6] Proteomics is important because it provides insight into the mechanisms and inner working of cells, which is critical to understand how diseases affect the host and to design appropriate targeted treatments. Bottom-up proteomics studies have

shown much success in the study of entire host proteomes over the years.[7] However, there are inherent disadvantages of bottom-up proteomics, including loss of PTMs,[8] inaccurate protein identification etc. Therefore, top-down (intact protein) proteomics studies have steadily increased in popularity in recent years for studying biological molecules/systems of interest. The top-down approach preserves the most information of the initial biological system, including intact protein sequences and chemical modifications (*e.g.* PTMs) as less disruptive chemicals are needed, which could alter/dissociate labile modifications and interactions in-between or within the proteins.

Drug target analysis is particularly important in drug design to enhance a drug's selectivity and specificity. However, one particularly challenging hurdle for top-down analysis using MS and MS/MS is the presence of disulfide bonds within proteins of interest. Cystine disulfide bonds are covalent linkages within protein structures and occur frequently in both intra-protein and inter-protein linkages. The formation of disulfide bonds in proteins is an important PTM related to protein folding and conferring stability on protein tertiary structures.[9,10] However, these disulfide linkages create cyclic regions within the protein sequence which hinder fragmentation and the production of sequence informative fragments, limiting the fragmentation efficiency upon MS/MS. Breaking disulfide linkages is often achieved by using chemical-based reducing reagents, such as dithiothreitol (DTT), reducing the sulphur-sulphur linkage to 2 free thiols (SH). Reformation of disulfide bonds is often quickly observed after reduction (< 1 h), therefore, alkylating reagents such as iodoacetamide (IAA) are often added to the reduced sample, preventing the reformation of disulfide bonds by reacting with the free thiols. These chemicals are often used in bottom-up or top-down analysis, however, these reactions often use very harsh reagents and conditions, *e.g.* using excess of reducing and alkylation reagents to maximise their effectiveness and turnover, which can easily damage certain protein/modification functional groups, and can easily modify analytes artificially providing misleading information during analysis and requiring further sample preparation/clean up steps downstream. Alternative reduction techniques have therefore been studied to overcome the use of chemical-based reduction.

Electrochemistry (EC) has been shown to reduce disulfide bonds successfully, enhancing the fragmentation performance and increasing the protein sequence cleavage coverage upon MS/MS.[11,12] By reducing the disulfide bonds electrochemically, no chemical-based reducing reagents are needed during sample preparation, meaning the protein structure/modifications are less likely to be affected, and no extra clean up steps are required prior to MS analysis. The electrochemical cell can also be hyphenated with MS for online disulfide bond reduction, allowing more effective reduction without extra sample preparation, preventing the loss of sample or contamination when performing the reduction process separately.

Plasma-induced or ultraviolet induced radicals have also been shown to be able to reduce disulfide linkages successfully. Xia *et al.* have shown plasma-induced radicals, often hydroxyl radicals, initiated from a low-temperature helium plasma were able to break disulfide bonds *via* dissociative addition.[10,13] Later, Xia *et al.* discovered that ultraviolet photo-induced radicals can also break disulfide linkages by irradiating the nano electrospray ionisation (nESI) plume with ultraviolet photons (185 and 254 nm). Small oxygen-containing radical species can be produced from oxygen in air and react *in situ* with the disulfide bonds within the nESI plume. Oxygen in air absorbs energy from ultraviolet photons and generate reactive atomic oxygen ($O(^3P)$) and metastable excited oxygen ($O(^1D)$), which further react with the water molecules in the solvent system and produce hydroxyl radicals (•OH). The hydroxyl radicals then react with the saturated alcohol (methanol, CH_3OH) in the solvent system and generate carbon-centred hydroxyalkyl radicals ($\bullet CH_2OH$, Scheme 3.1).[14] The hydroxyalkyl radicals have been shown to break the disulfide bonds and allow better cleavage coverage for tandem mass spectrometry experiments.



Scheme 3.1 (a) oxygen in air irradiated by UV photons and becomes reactive atomic oxygen, (b) metastable excited oxygen reacts with water and create hydroxyl radicals, (c) Hydroxyl radicals react with methanol and generate hydroxyalkyl radicals and water.[14]

Top-down analysis often involves larger proteins, therefore, extremely accurate measurement of the protein's intact mass-to-charge ratio (*m/z*) and MS/MS fragments are critical for top-down proteomics analysis, to resolve and identify fragments, PTMs, and thus, proteins and proteoforms. Fourier transform-ion cyclotron resonance mass spectrometry (FT-ICR MS) is an analytical technique especially well-known for its ultra-high resolution, routinely 500,000 - 1,000,000 resolving power,[15] and the highest mass accuracy, usually obtaining sub parts-per-million (ppm) to parts-per-billion (ppb) ranges of mass errors.[16] FT-ICR MS also offers the widest array of fragmentation techniques, such as collisionally activated dissociation (CAD),[17] electron capture dissociation (ECD),[18] electron induced dissociation (EID),[19] electron detachment dissociation (EDD),[20] infrared multiphoton dissociation (IRMPD),[21] ultraviolet photodissociation (UVPD),[22] etc. These important features offered by FT-ICR MS allow extremely accurate measurement of the protein's intact *m/z* and provide sequence-informative fragments using various fragmentation techniques.

Herein two different instrumentation-based methods are investigated; electrochemical reduction and radical-induced reduction of disulfide bonds within protein structures, enabling more effective MS and MS/MS experiments.[23] Both

techniques enabled a large increase in subsequent MS/MS fragmentation and sequence cleavage coverage. In addition, analysis of isotopic patterns shows that the EC reduction process occurs in a sequential manner depending on the electrode potential used, enabling selective and sequential reduction of disulfide bonds within complex protein architectures.

Experimental

Materials

Human lysozyme (hLZ), human insulin, bovine ribonuclease A (RNAase A), formic acid, dithiothreitol (DTT), and iodoacetamide (IAA) were all purchased from Sigma Aldrich (Dorset, England). LCMS grade acetonitrile (MeCN) was purchased from VWR (Lutterworth, UK). LCMS grade methanol (CH_3OH) was purchased from Thermo Fisher Scientific (Waltham, UK). Ultra-pure water (18.2 $M\Omega cm^{-1}$) was obtained from a Milli-Q UV III water purification system (Milli-Q, Hertfordshire, UK).

Electrochemical Reduction

Electrochemical reduction of disulfide bonds was achieved using a ROXY EC system (Antec, The Netherlands), consisting of a potentiostat, equipped with a µ-PrepCell 2.0 fitted with a proprietary Ti-Blue electrode (methylene blue coated titanium electrode), and a syringe pump. Samples (non-reduced intact proteins) were passed through an electrochemical cell, at a constant flow rate ~ 1-2 µL/minute, using a syringe pump (Figure 1). Disulfide bonds present in the sample were reduced when the sample was passed through the electrochemical cell. The potentiostat provided a consistent negative voltage for disulfide bond reduction. The µ-PrepCell 2.0 is a thin-layer electrochemical cell consisting of a Ti-Blue working electrode (WE), a conductive polymeric auxiliary electrode (conductive PEEK) (AUX), and a palladium/hydrogen (Pd/H_2) reference electrode (REF). A 75 µm spacer was used between the AUX electrode and base EC cell to provide a flow path, allowing smaller dead volumes than previous 150 µm spacer designs. The flow rate and spacer height dictated dwell time within the cell, thus affecting reaction time and products. These were optimised to be balanced with good ESI spray conditions to 75 µm spacer and 1-2 µL/minute flow rate.

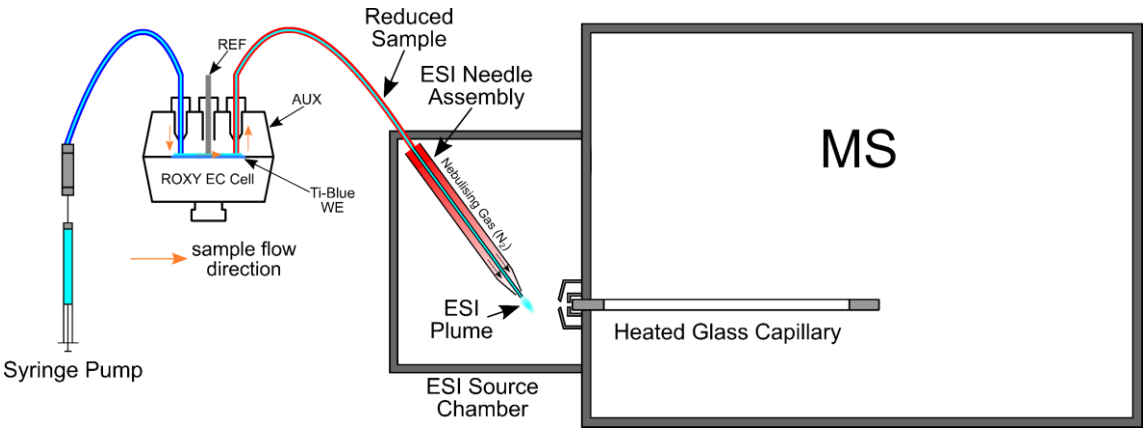


Figure 3.1 Schematic representation of the electrochemical disulfide bond reduction set up. A ROXY electrochemical cell is installed after the syringe pump and prior to the mass spectrometer. The samples were reduced in the ROXY electrochemical cell (orange arrows indicate sample flow direction). REF is the reference electrode, AUX is the auxiliary electrode, and WE is the working electrode.

Radical-Induced Reduction

A 2.54 cm lamp length, 0.64 cm diameter, double bore tubing low pressure mercury (LP Hg) lamp with a synthetic quartz lamp case (model number: 81-1057-51; BHK, Inc., California, USA), operating at 20 mA, was used to initiate the radical reactions by generating hydroxyalkyl radicals. The lamp specifications from the manufacturer states the minimum intensity with the primary wavelength (254 nm) at 20 cm distance was measured 73 µW/cm^2 and 42 µW/cm^2 for 185 nm measured in an argon atmosphere. The mercury lamp was placed ~ 2 cm away from the centre of the nESI plume (Figure 2). The arc lamp was driven by a power supply (start voltage of 1050 V and operational voltage of 320 V) (BHK, Inc., California, USA) which allowed stable use for many hours with no noticeable change.

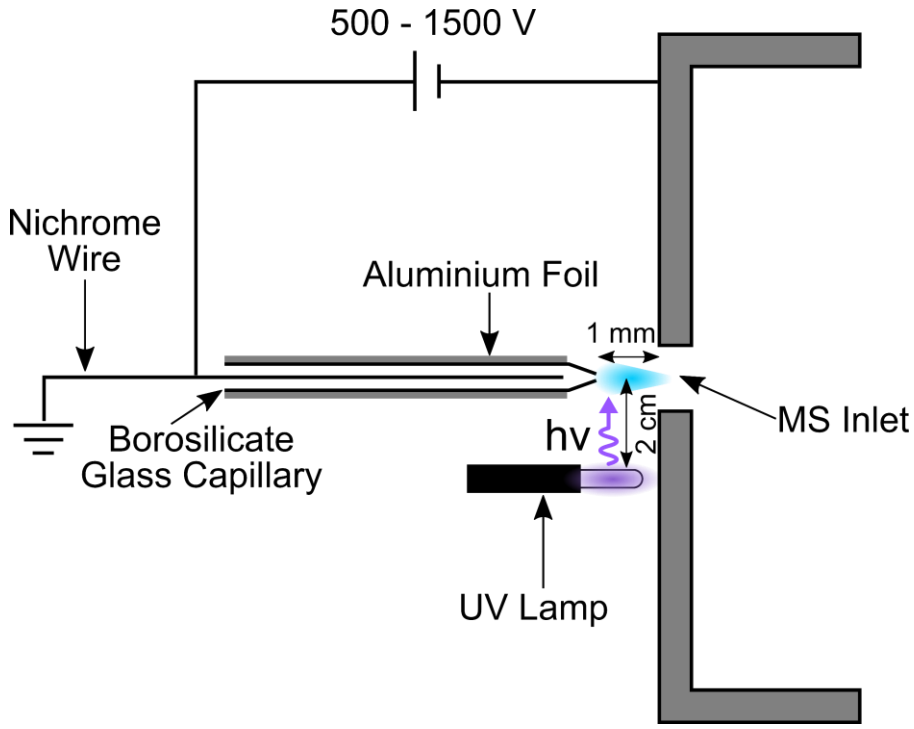


Figure 3.2 Schematic representation of the radical-induced disulfide bond reduction set up. The UV lamp was placed 2 cm next to the nESI plume, allowing the UV photons to react with the methanol in the sample and create hydroxyalkyl radicals.

Mass Spectrometry

Samples for electrochemical reduction were analysed with electrospray ionisation (ESI) to achieve a constant flow rate of ~ 1-2 µL/minute. Samples with radical-induced reduction were analysed with nESI using borosilicate glass capillary, pulled with a P-97 Flaming/Brown micropipette puller (Sutter Instrument, California, USA). All experiments were carried out on a SolariX FT-ICR mass spectrometer, fitted with a 12 tesla actively shielded magnet (Bruker Daltonik GmbH, Bremen, Germany). All samples were sprayed in positive ion mode. For MS experiments, ions were accumulated for 0.1-0.5 s in the hexapole-based collision cell before transferred to the infinity cell for excitation and detection.[24]

For CAD MS/MS experiments, ions of interest were isolated using the front-end quadrupole using an isolation window of 10-30 *m/z*. Selected ions were then accelerated

into argon collision gas at 18-25 V. Ions were continuously accumulated for 3 s before transmission and detection.

For ECD MS/MS experiments, ions of interest were isolated using the front-end quadrupole using a window of 4-10 *m/z*, then accumulated in the hexapole for 3-4 s. Isolated ions of interest were then transferred and trapped in the infinity cell. Trapped ions were then irradiated for 15-25 ms with 1.2 eV electrons produced from a 1.5 A hollow cathode dispenser *via* an extraction lens held at 18 V.

All spectra were internally calibrated using a quadratic calibration function[25] and peak picked using Sophisticated Numerical Annotation Procedure (SNAP) algorithm v2.0 using the following parameters: quality factor of 0.80, a signal-to-noise (S/N) threshold of 2, and a maximum charge state of 20+, a constant building block (base elementary composition) of C = 4.9384, N = 1.3577, O = 1.4773, S = 0.0417, H = 7.7583. All spectra were then manually interpreted and assigned *via* DataAnalysis v4.2 (Bruker Daltonik GmbH, Bremen, Germany).

Results and Discussion

Electrochemical Reduction

Heavily disulfide-bonded proteins (such as lysozyme (4 disulfides), RNAase A (4 disulfides), and antibodies (16 disulfides) etc.) were ionised and analysed using ESI-MS, but are often observed at lower charger states (higher *m/z*) and narrower charge state distributions compared to non-disulfide-bonded analogous species as the protein structures are held together with cystine linkages, preventing further regions being protonated. As a result of these effects, observation of high charge states for lower limit of detection and for more extensive fragmentation in MS/MS experiments is often limited. Thus, the use of a denaturing/acidified solvent system has become a common technique to maximise the level of protonation on biomolecules by unfolding the protein structure. However, due to the presence of these cystine linkages, even in a denaturing/acidified solvent system, many proteins, such as human lysozyme (hLZ) and ribonuclease A (RNAase A) were observed in relatively low charge states, from 8+ to 12+ (Figure 3.3a and 3.4a respectively). However, with electrochemical reduction,

whole new arrays of charge state distributions were observed for both hLZ and RNAase A samples when disulfide bonds were reduced at -2.0 V with the ROXY-EC system. A charge state distribution of 8+ to 20+ (8 new charge states) was observed for hLZ (Figure 3.3b) and 8+ to 18+ (6 new charge states) for RNAase A (Figure 3.4b).

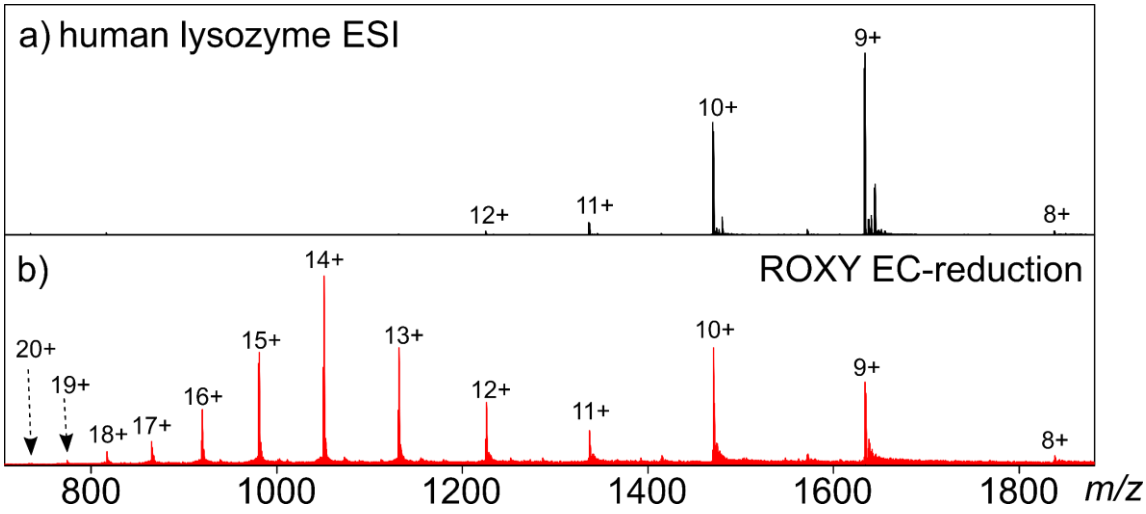


Figure 3.3 ESI-MS spectra of hLZ under denaturing conditions (50:50 H_2O:MeCN + 1% formic acid). a) no reduction by ROXY-EC system, b) reduced by ROXY-EC system at -2.0 V.

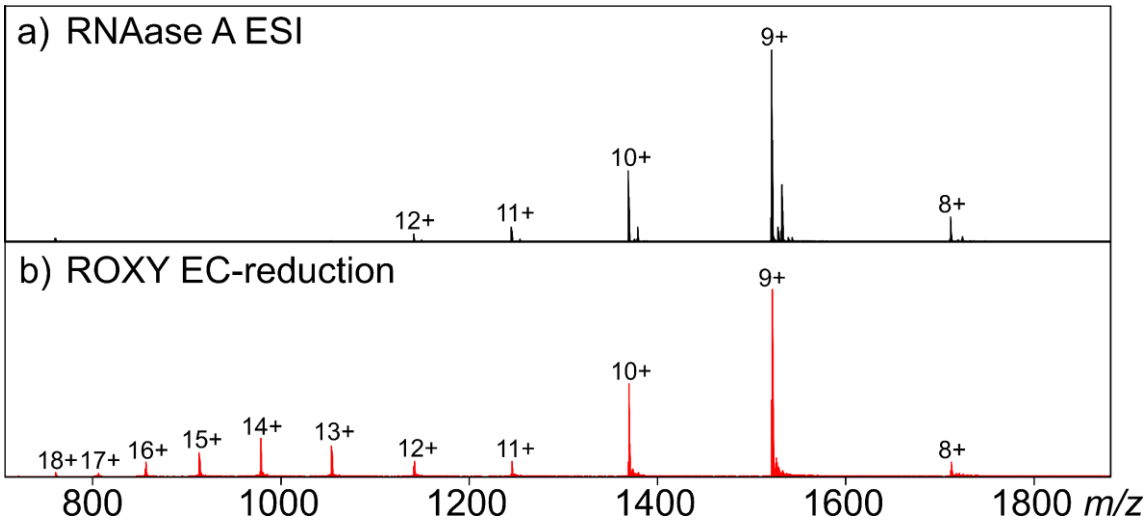


Figure 3.4 ESI-MS spectra of RNAase A under denaturing conditions (50:50 H_2O:MeCN + 0.1% formic acid). a) no reduction by ROXY-EC system, b) reduced by ROXY-EC system at -2.0 V.

For species such as insulin[26] and antibodies[27], the reduction of disulfide bonds results in the separation of peptide chains, forming distinctly different product species, and be observed in the resulting mass spectrum. However, for species having intramolecular disulfide bonds such as hLZ and RNAase A, reduction of each disulfide bond only results in an addition of two nominal mass units (2H, 2.015650 Da). Due to the extended isotopic pattern of larger biomolecule species, separating the new reduced species (+2H) isotopic pattern from that of the non-reduced species is particularly challenging, even for ultra-high resolution instrumentation such as FT-ICR MS (Figure 3.5).

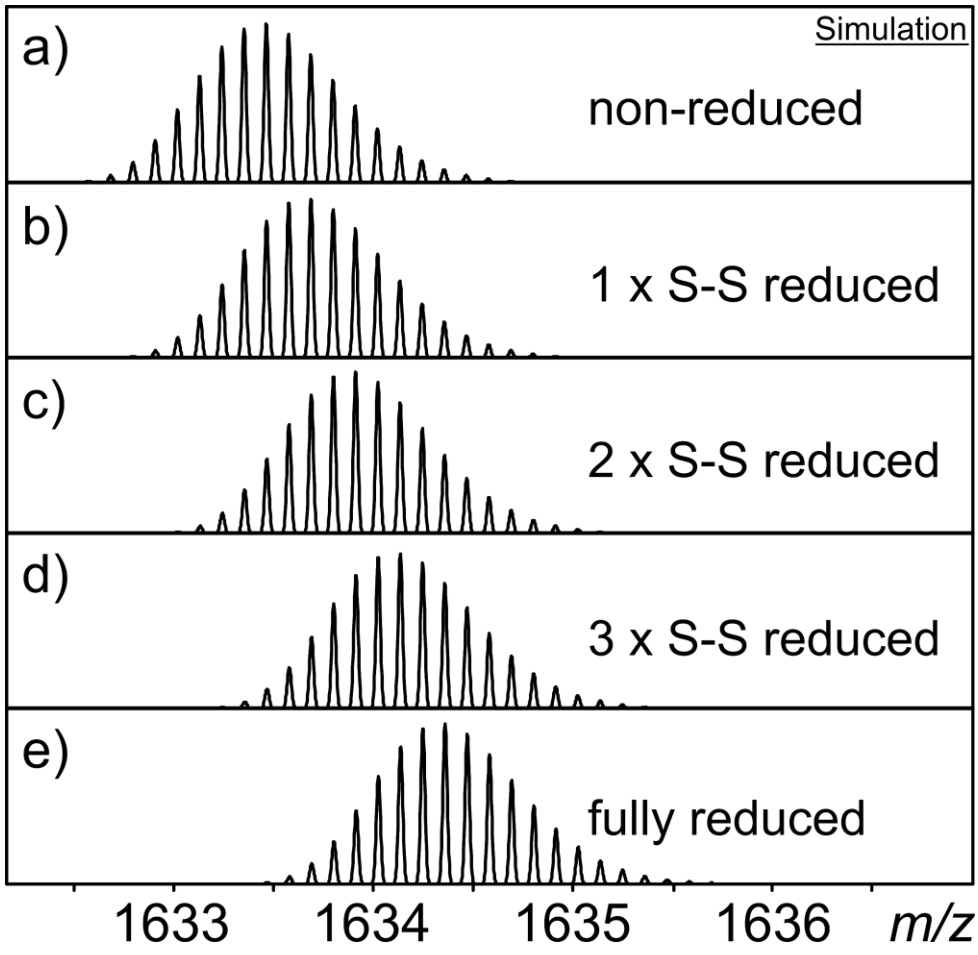


Figure 3.5 Isotopic simulations of the reduction products of hLZ (charge state 9+). a) non-reduced hLZ, b) hLZ with 1 disulfide bond reduced, c) hLZ with 2 disulfide bonds reduced, d) hLZ with 3 disulfide bonds reduced, e) hLZ with 4 disulfide bonds reduced (fully reduced).

Instead, a statistical approach was adopted in which each individual isotope and its peak area was measured from each charge state observed for each reduction potential (14 different potentials, from -2.6 V to 0 V, with increments of 0.2 V) produced from the ROXY-EC system, constituting ~ 2500 isotopes measured for each protein (hLZ and RNAase A) At certain reduction potentials, some disulfide bonds were reduced, while others remained, the isotopes could then be used to calculate how many disulfide bonds had been reduced by comparing the isotope peak areas with that of the non-reduced isotopic pattern. Isotopic peak measurements were weight-averaged and scaled according to their differences between the non-reduced protein form (*e.g.* Figure 3.5a for hLZ) and the fully reduced protein form (Figure 3.5e for hLZ). This scaling provided a percentage of reduction measurement, which could easily be converted to the number of disulfide bonds *via* division by the number of disulfide bonds in the protein of interest (*e.g.* the 4 disulfide bonds in hLZ equate to 25% reduction per disulfide bond). The percentage of reduction measurement was calculated with the equations below.

$$\overline{x_i} = \frac{\sum_i[(m/z)^2 A]}{\sum_i[(m/z) A]} \qquad \text{[Eq. 1]}$$

where $\overline{x}_i$ is the weighted average of the i+ charge state at the reducing potential, *m/z* is the mass-to-charge ratio of individual isotope in the i+ charge state distribution, and A is the corresponding peak area of the isotope.

$$scaled\ \overline{x_i} = \frac{\overline{x_i} - \overline{x_i}_{(non-reduced)}}{\overline{x_i}_{(fully\ reduced)} - \overline{x_i}_{(non-reduced)}} \times 100\% \qquad \text{[Eq. 2]}$$

where scaled $\overline{x}_i$ is scaled weighted average of the i+ charge state at the reducing potential, $\overline{x}_{i(\text{non-reduced})}$ is the weighted average of the i+ charge state from simulation, and $\overline{x}_{i(\text{fully reduced})}$ is the weighted average of the i+ charge state from simulation.

$$average\ number\ of\ S-S\ reduced = \frac{\sum_i[(scaled\ \overline{x_i})^2 \left(\frac{\sum A_i}{\sum A_{all}}\right)]}{\sum_i[(scaled\ \overline{x_i}) \left(\frac{\sum A_i}{\sum A_{all}}\right)]} \qquad \text{[Eq. 3]}$$

where ΣA_i is the sum of peak areas of all isotopes from i+ charge state, and ΣA_{all} is the sum of peak areas of all isotopes from all charge states.

The results (shown below in Figures 3.6 and 3.7) show that the electrochemical reduction process is sequential, and that disulfide bonds can be selectively reduced by applying the correct potential. The results also show that as disulfide bonds are reduced, they are able to unfold even more under the denaturing conditions used (50:50 H_2O:MeCN + 1% formic acid) and attain higher charge states during the ESI process.

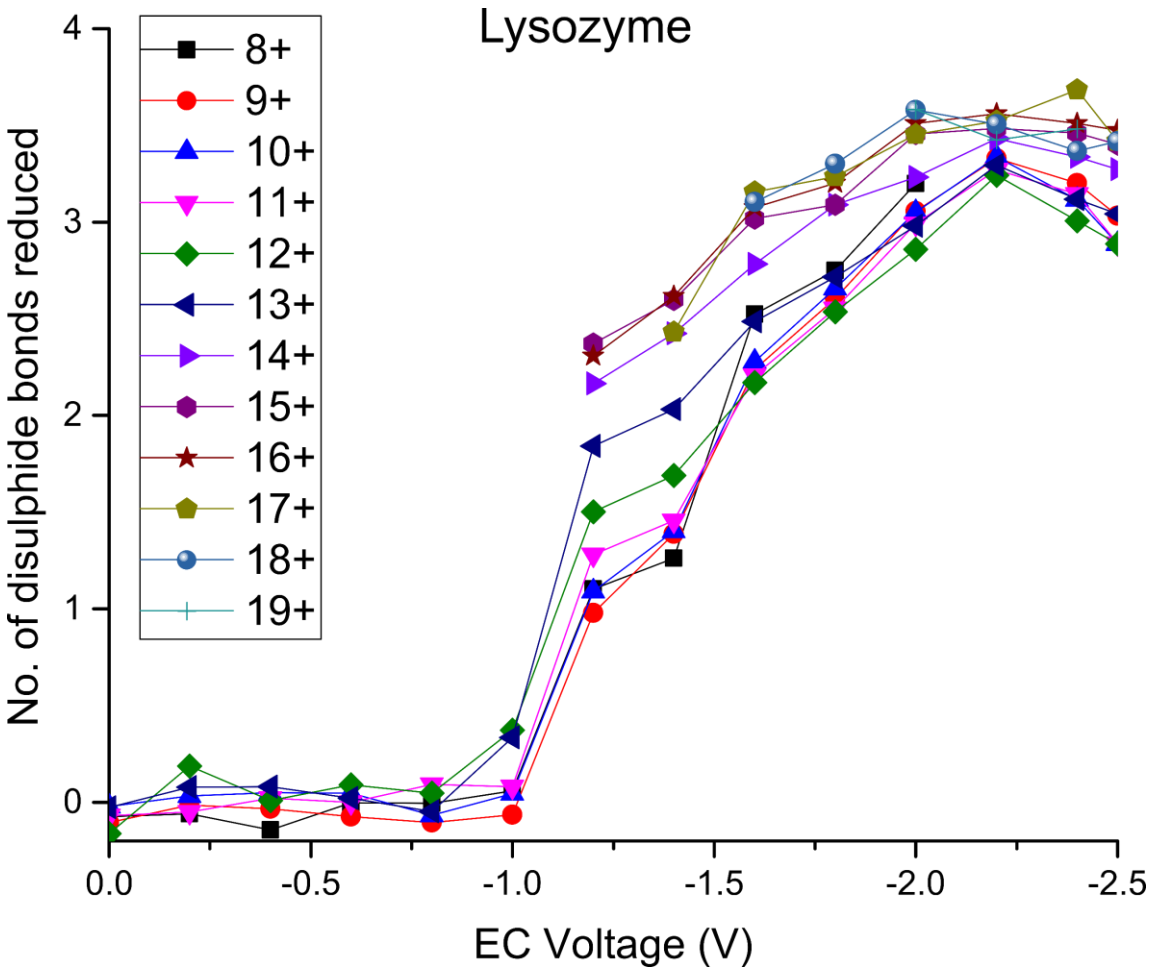


Figure 3.6 Number of disulfide bonds reduced as a function of EC potential using the ROXY EC system to reduce hLZ. Each charge state is shown separately to show transitions of proteins to higher charge states upon reduction.

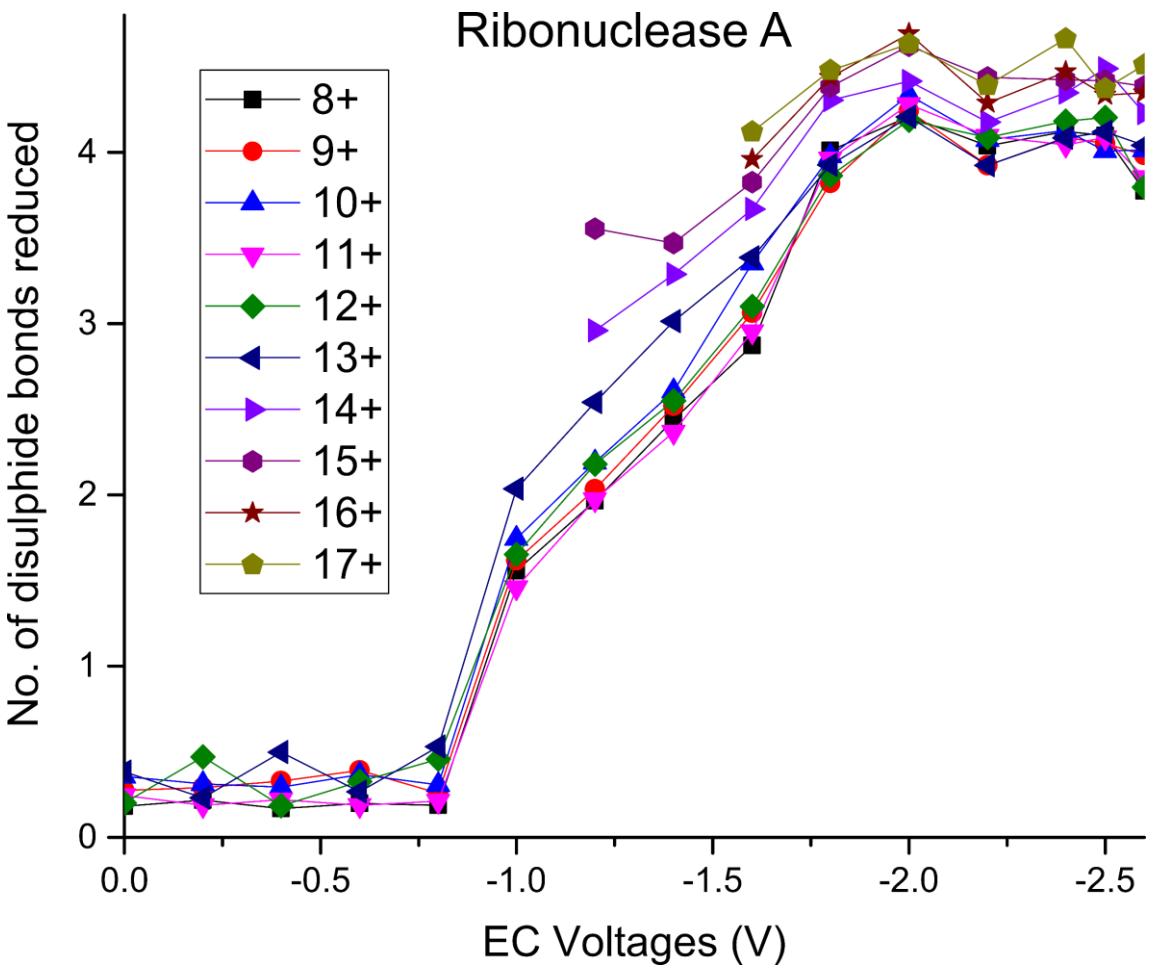


Figure 3.7 Number of disulfide bonds reduced as a function of EC potential using the ROXY EC system to reduce RNAase A. Each charge state is shown separately to show transitions of proteins to higher charge states upon reduction.

Figures 3.6 and 3.7 show that the more reduced protein species occupy most of the higher charge states, while the less reduced species still reside within the lower charge states more naturally seen from the control spectra (Figures 3.3a and 3.4a). This approach also demonstrated the inaccuracy with using one charge state to monitor the reduction process, as reduced (and thus more unfolded) species readily adopt more highly charged states, monitoring a single, lower charge state from the control spectrum would lead to an underestimate of the overall reduction taking place.

Full reduction of all 4 disulfide bonds in each protein can be observed in both hLZ and RNAase A. The hLZ protein species was calculated to be fully reduced at a

potential of -2.2 V (at certain charge states), while RNAase A at -2.0V. The average number of disulfide bonds reduced at a given potential, using the findings shown in Figures 3.6 and 3.7, are shown in Figure 3.8, emphasising the differences for the two proteins studied.

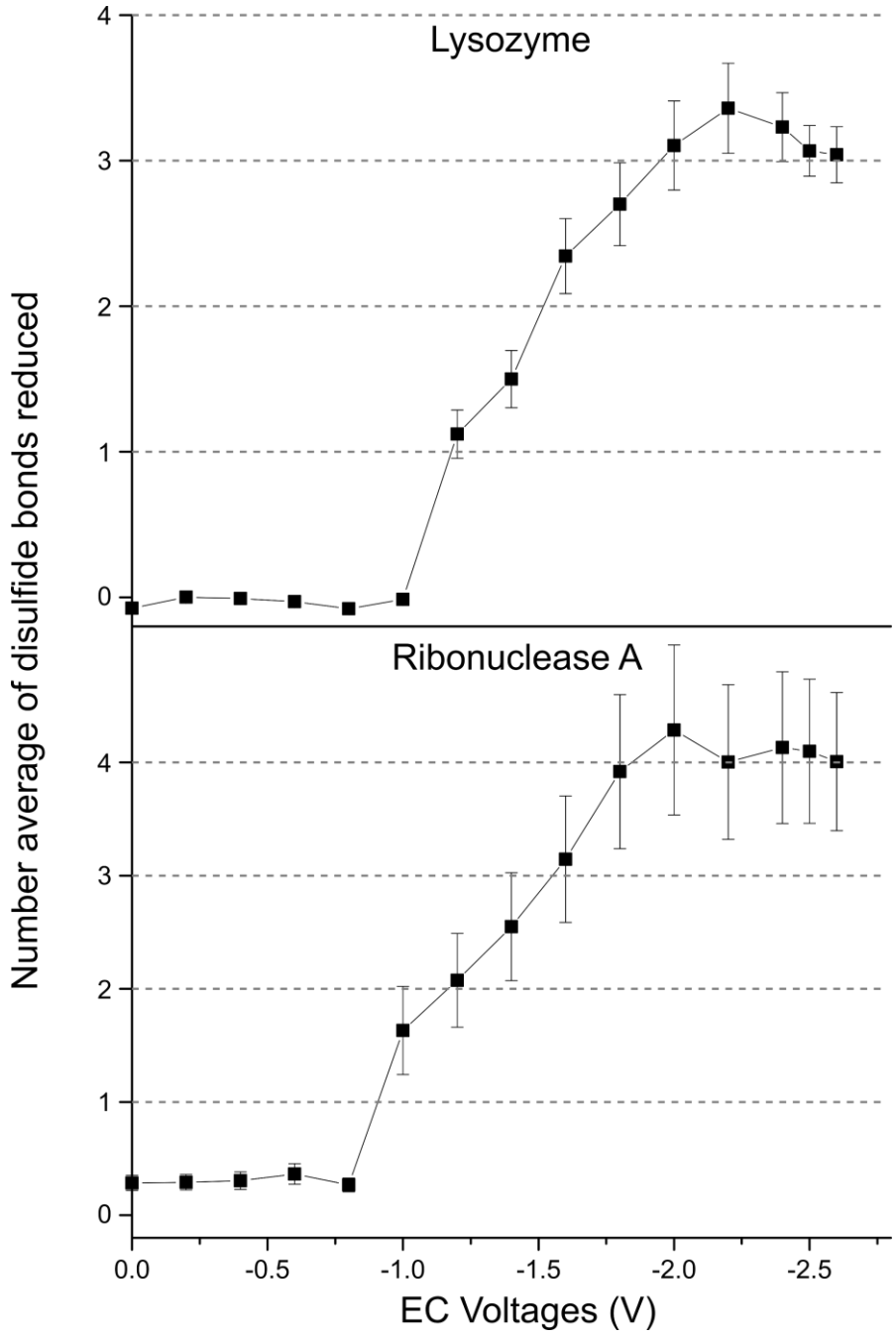


Figure 3.8 Average number of disulfide bonds reduced at each potential using the ROXY-EC system for hLZ (top) and RNAase A (bottom).

Human insulin was also used to investigate the reduction performance of the ROXY-EC system (Figure 3.9). Unlike hLZ and RNAase A, when insulin is fully reduced, the two polypeptide chains (Chain A and Chain B) will be dissociated, which can be used to monitor the reduction event. Using ESI-MS, 4+ to 6+ charge states of intact insulin were observed ($[M+nH]^{n+}$) Separated chains (Chain A (C_A) and Chain B(C_B)) were then observed with increasing reduction time. The longer the reaction time, the higher the percentage of conversion (from intact insulin to the separated C_A and C_B) (Figure 10). The percentage of conversion can be calculated using the equation below.

$$Percentage\ of\ Conversion = \frac{\Sigma(A_{(C_A+C_B)})}{\Sigma(A_{(intact)})} \times 100\% \qquad \text{[Eq. 4]}$$

where $\Sigma(A_{(CA+CB)})$ is the sum of peak areas of C_A and C_B, $\Sigma(A_{(intact)})$ is the sum of peak areas of intact insulin.

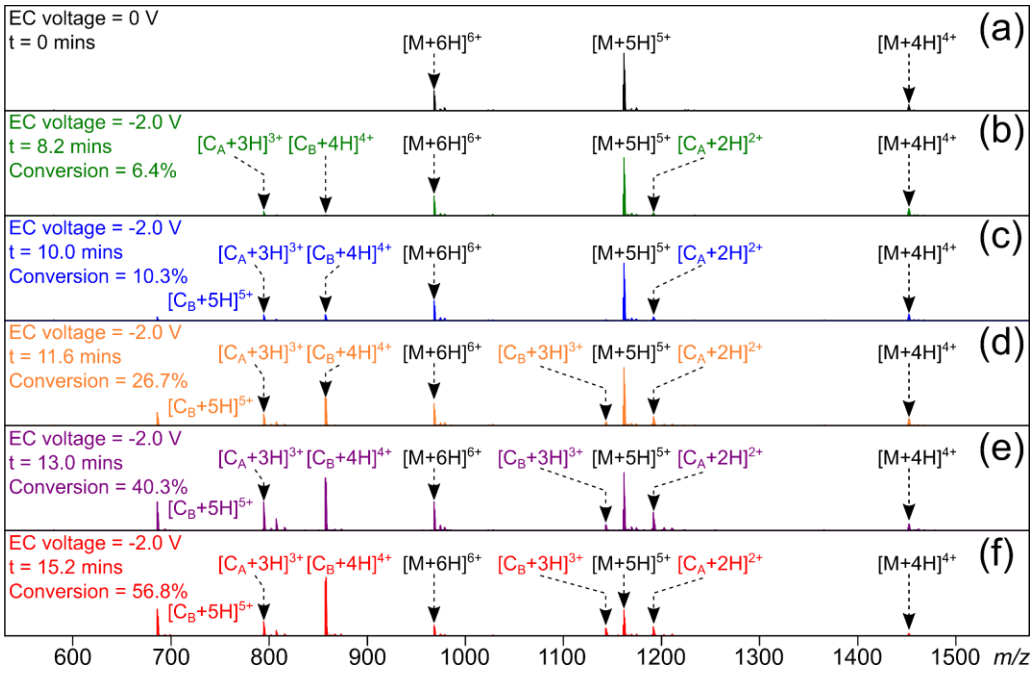


Figure 3.9 ESI-MS of insulin upon electrochemical reduction (EC voltage = -2.0 V) over a time period of 15.2 min. a) t = 0 and no reduction, b) t = 8.2 min and conversion of 6.4%, c) t = 10.0 min and conversion of 10.3%, d) t = 11.6 min and conversion of 26.7%, e) t = 13.0 min and conversion of 40.3, and f) t = 15.2 min and conversion of 56.8%.

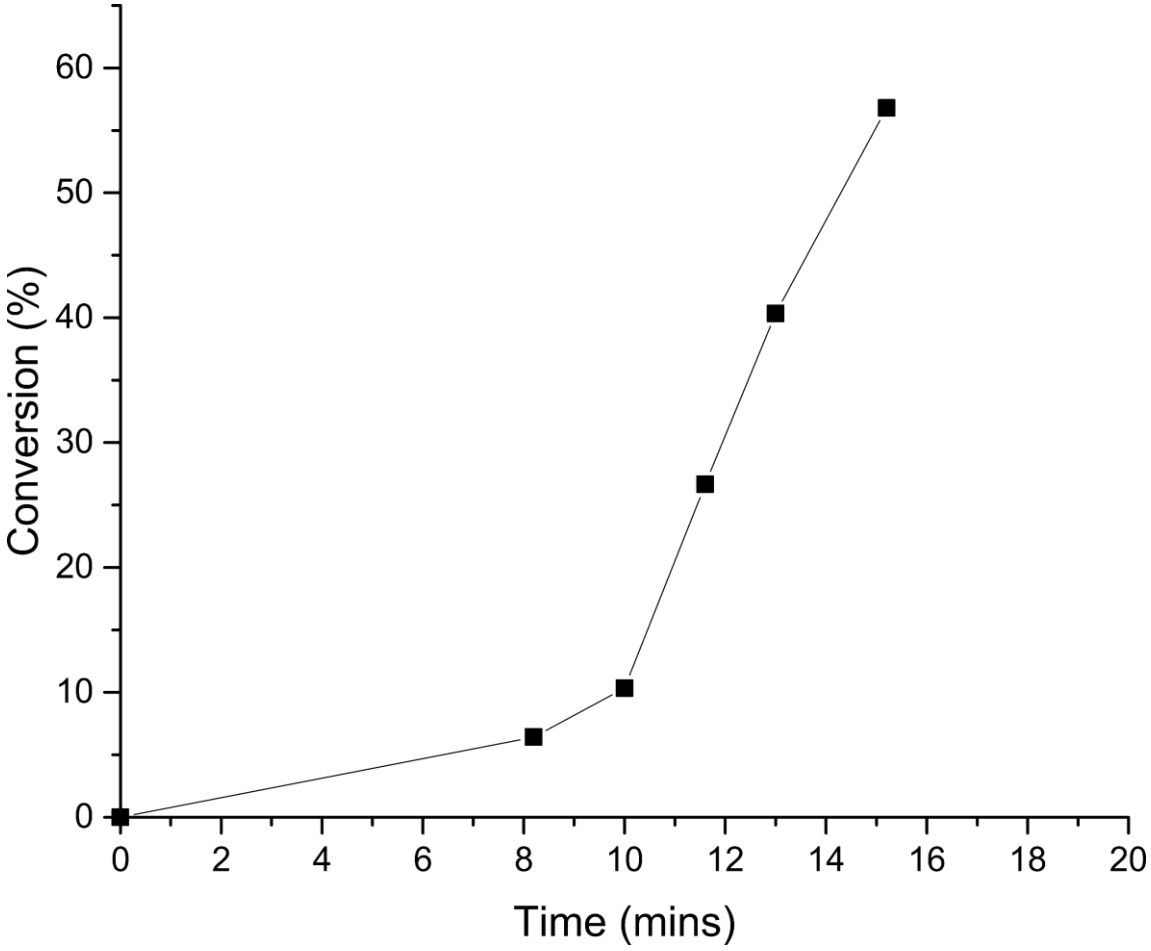


Figure 3.10 Plot of percentage conversion (from intact to C_A and C_B) as a function of time.

Higher charge states have been shown to provide more extensive fragmentation during MS/MS experiments, and thus improve sequence/cleavage coverage, and overall protein characterisation.[23] These benefits were also highlighted for the species observed above. The hLZ species detected were subjected to CAD MS/MS and ECD MS/MS fragmentation in both their non-reduced and reduced forms. Figures 3.11 and 3.12 (Table S3.1) show the CAD MS/MS spectrum and the fragmentation map for the non-reduced hLZ species. From the fragmentation map in Figure 3.12, fragments near the termini are observed, as expected with CAD MS/MS, and it is clearly shown that the "disulfide bond protected" regions were not fragmented, yielding a cleavage coverage of 10%. However, by applying an EC voltage of -2.2 V, the disulfide linkages were reduced and allowed more protonation on the protein backbone, yielded an observation

of higher charge states (as shown in Figure 3.3b). By applying CAD MS/MS to the reduced hLC species, a slight increase of cleavage coverage was observed from 10% to 17% and more importantly, fragments were observed in the "disulfide bond protected" regions, allowing more sequence informative coverage (Figures 3.13 and 3.14, Table S3.2).

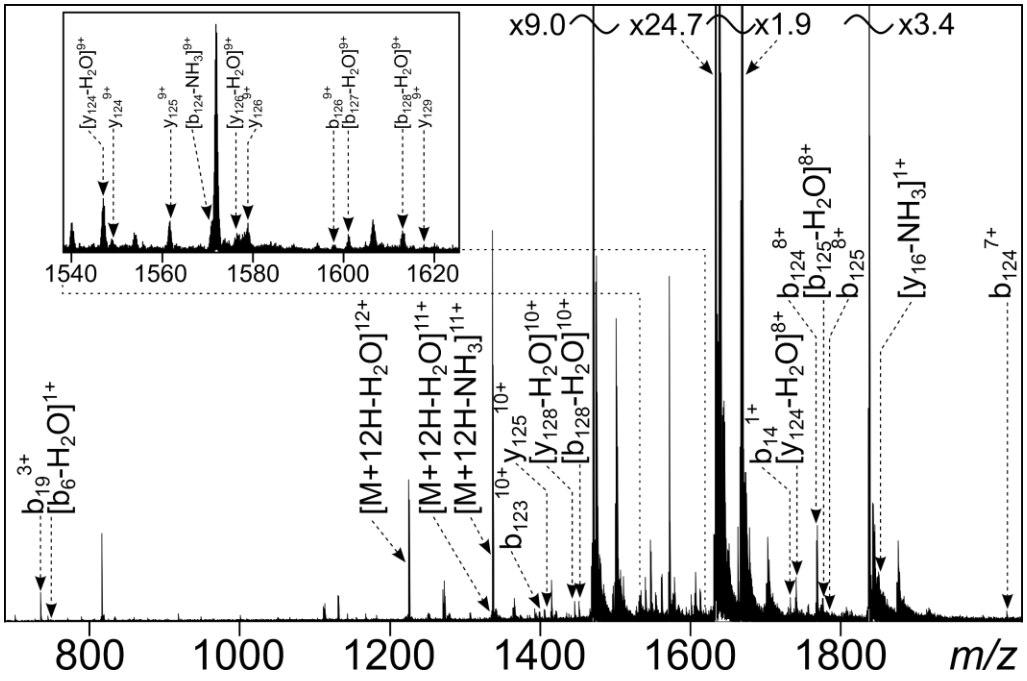


Figure 3.11 CAD MS/MS spectrum of the non-reduced (no EC reduction) hLZ species.

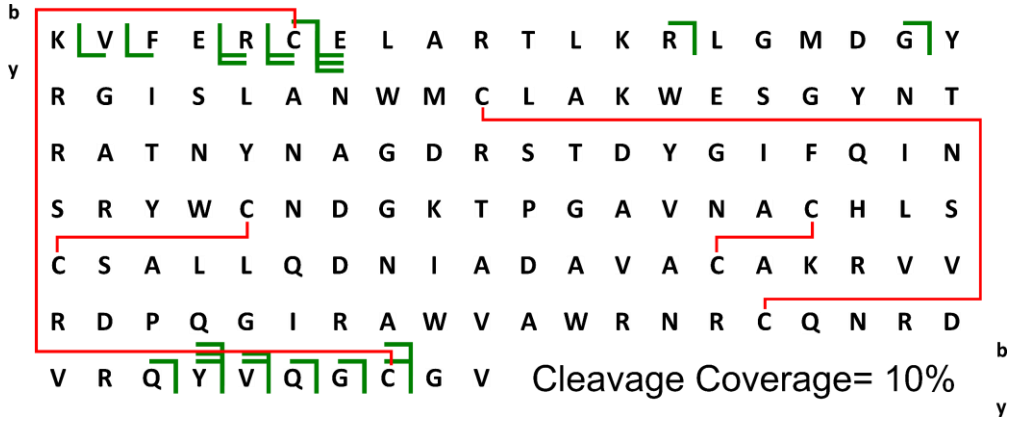


Figure 3.12 Fragmentation map for the CAD MS/MS spectrum of the non-reduced (no EC reduction) hLZ species, disulfide bonds are shown in red.

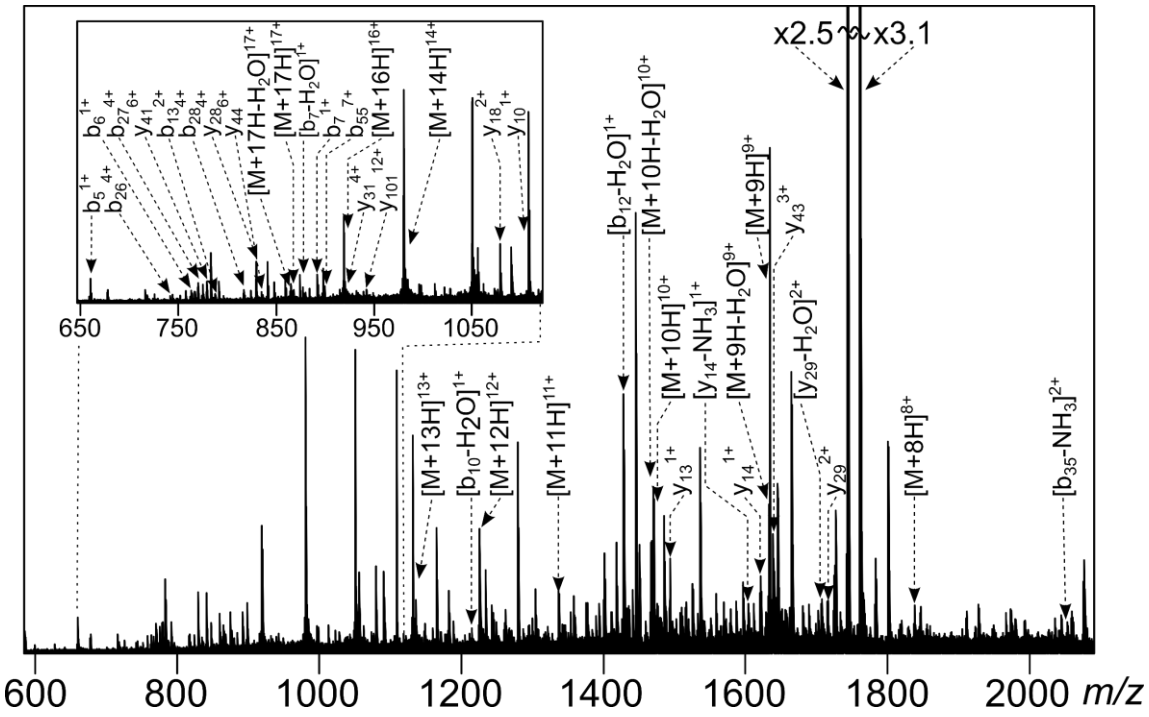


Figure 3.13 CAD MS/MS spectrum of the reduced hLZ species (EC voltage = -2.2 V).

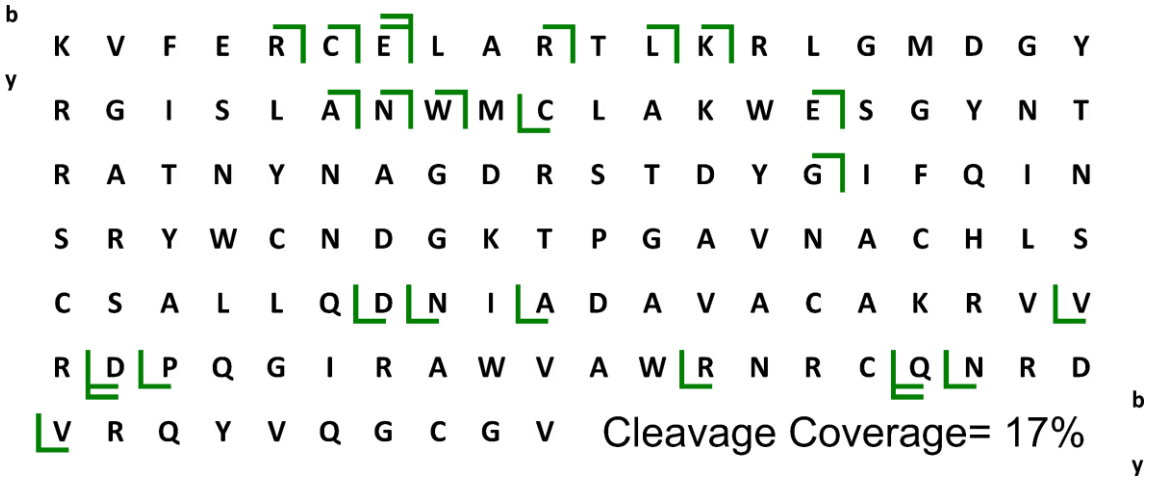


Figure 3.14 Fragmentation map for the CAD MS/MS spectrum of the reduced hLZ species (EC voltage = -2.2 V).

ECD MS/MS was also applied to investigate the difference in cleavage coverage between the non-reduced and reduced hLZ species. Figures 3.15 and 3.16 (Table S3.3) show the ECD MS/MS spectrum and the fragmentation map of the non-reduced hLZ species. Like CAD MS/MS of the same species, there were minimal fragments

observed in the "disulfide bond protected" regions, yielding a total cleavage coverage of 3%, even though electrons were successfully captured by the non-reduced species, shown by the observation of a series of charge reduced species (CRS). By applying an EC voltage of -2.2 V, the hLZ species was fully reduced and upon ECD MS/MS, a vast increase of 22% in cleavage coverage was observed (from 3% to 25%, ~7-fold increase), allowed much more extensive fragmentation and characterisation of the protein species, including the "disulfide bond protected" regions (Figures 3.17 and 3.18, Table S3.4).

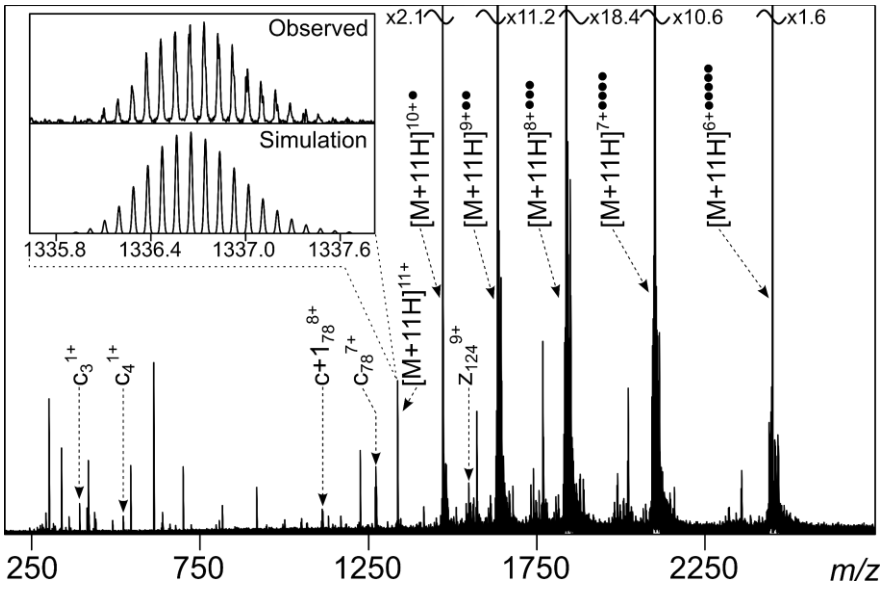


Figure 3.15 ECD MS/MS spectrum of the non-reduced (no EC reduction) hLZ species.

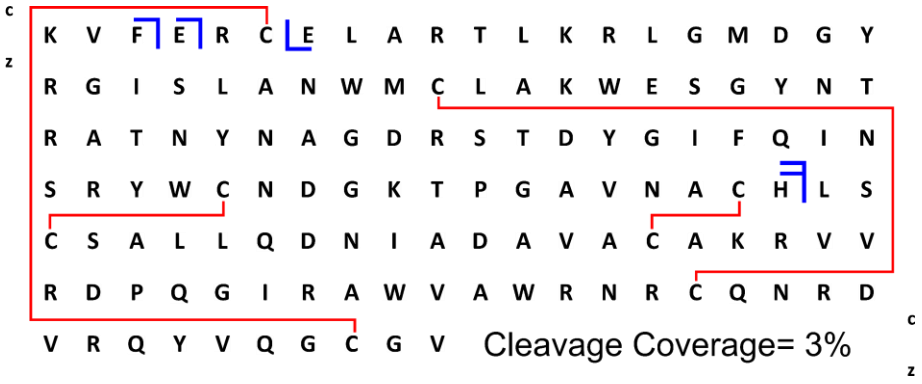


Figure 3.16 Fragmentation map for the ECD MS/MS spectrum of the non-reduced (no EC reduction) hLZ species, disulfide bonds are shown in red.

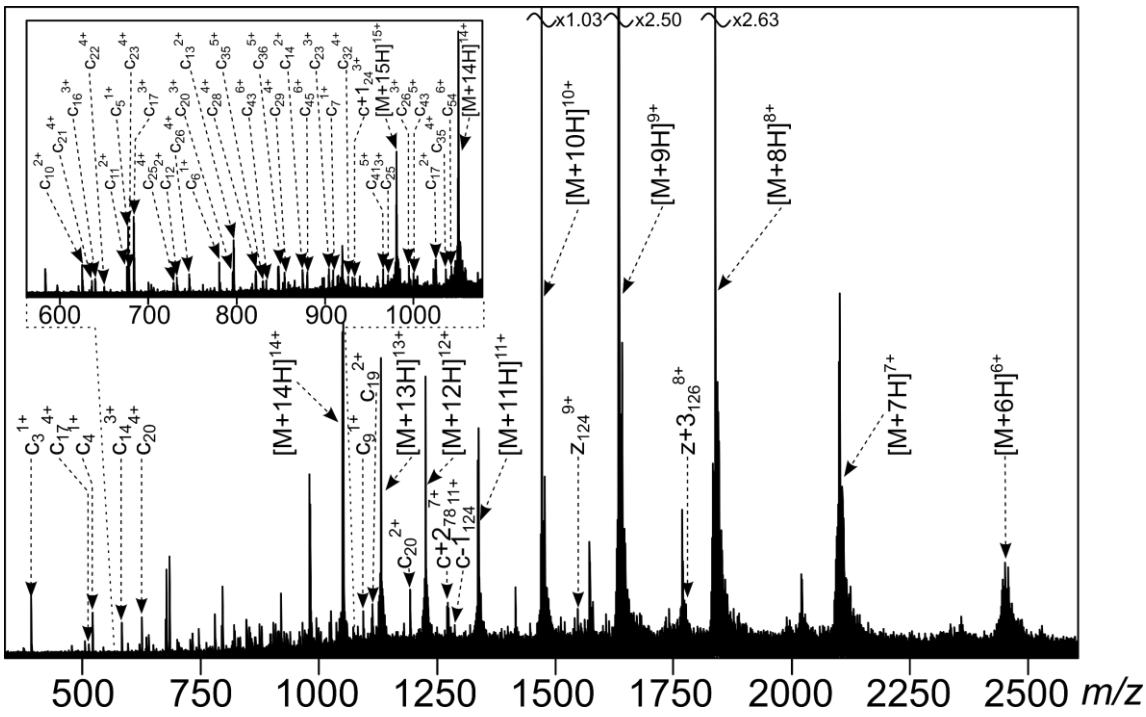


Figure 3.17 ECD MS/MS spectrum of the reduced hLZ species (EC voltage = -2.2 V).

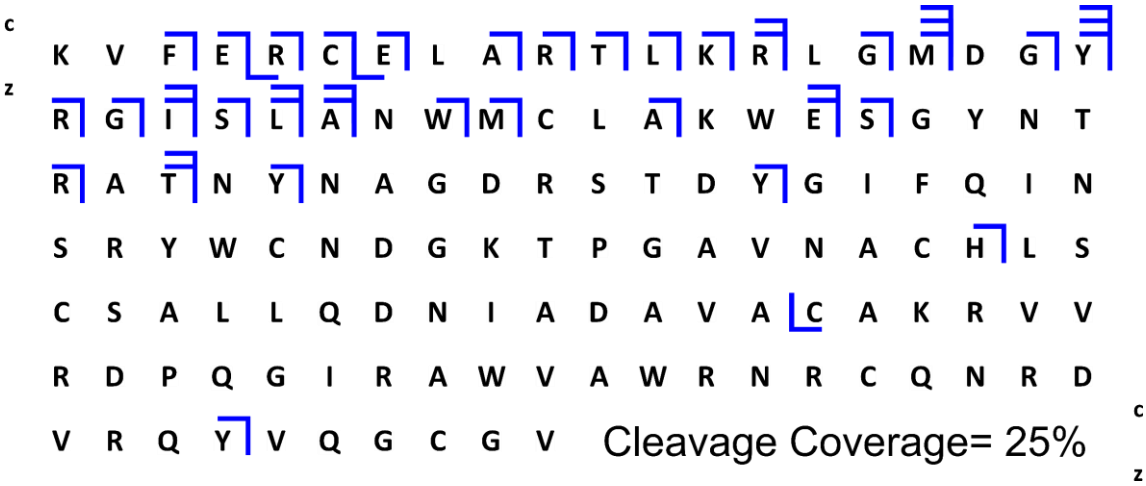


Figure 3.18 Fragmentation map for the ECD MS/MS spectrum of the reduced hLZ species (EC voltage = -2.2 V).

By comparing the cleavage coverage obtained by CAD and ECD MS/MS experiments on the non-reduced (EC voltage = 0V) and the fully reduced (EC voltage = -2.2 V) hLZ, a significant increase of 23% was observed (from 12% to 35%) (Figures 3.19 and 3.20). The increase in cleavage coverage on the protein backbone, even within

in “disulfide bond protected” regions, suggests the protein structure was opened up successfully upon EC reduction, allowing more protonation, and more extensive fragmentation.

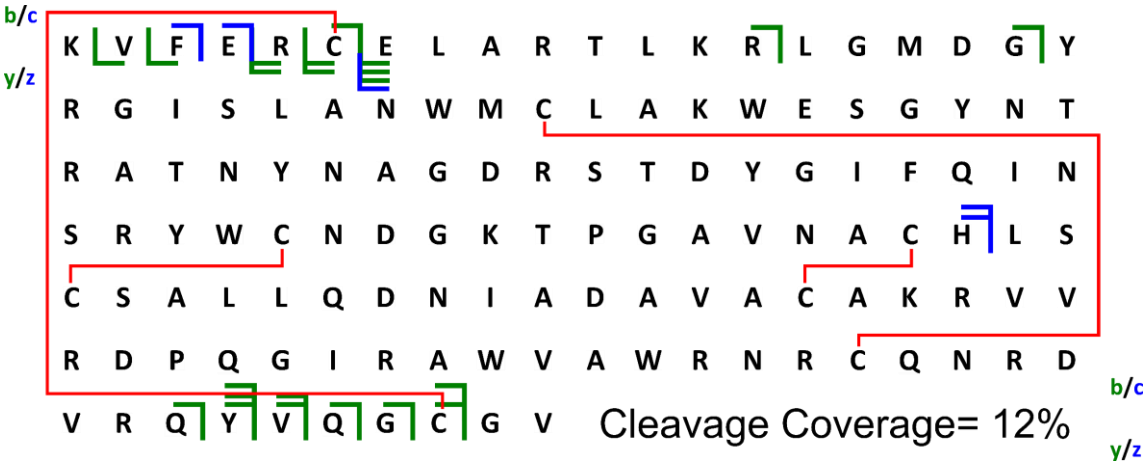


Figure 3.19 Combined fragmentation map of CAD and ECD MS/MS for the non-reduced hLZ (no EC reduction). b and y fragments are shown in green; c and z fragments are shown in blue. Disulfide bonds are shown in red.

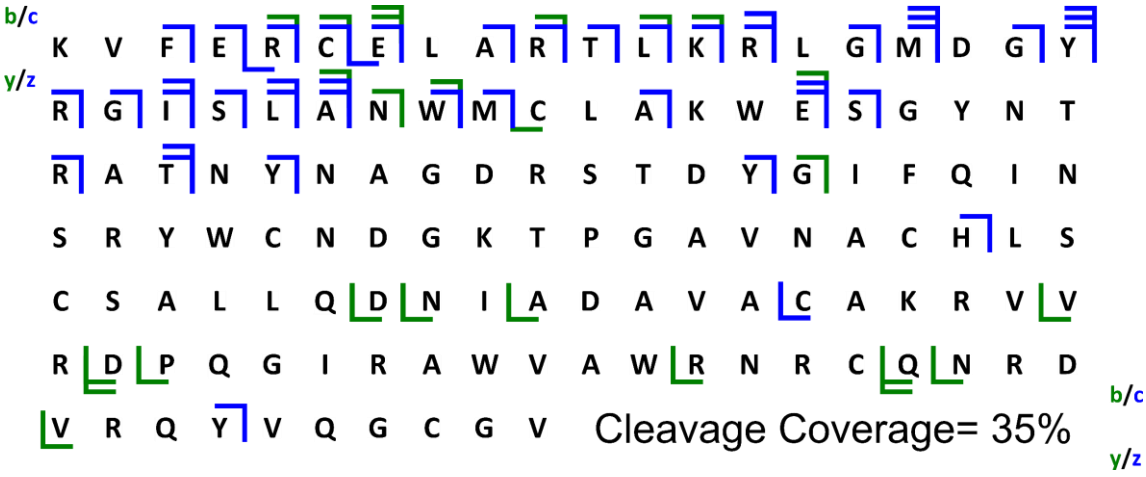


Figure 3.20 Combined fragmentation map of CAD and ECD MS/MS for the fully reduced hLZ (EC voltage = -2.2 V). b and y fragments are shown in green; c and z fragments are shown in blue.

Radical-Induced Reduction

Xia *et al.* previously reported the use of a mercury lamp to cleave disulfide bonds in model peptides, *e.g.* C*IELLQARC* (* indicates cysteine linkages), and achieved more fragments upon fragmentation.[14] By using a 50:50 H_2O:CH_3OH solvent system, and by irradiating with a mercury lamp, the UV photons can initiate a radical based reaction (Scheme 3.1), producing hydroxymethyl radicals ($\bullet CH_2OH$), which can react with free thiols preventing disulfide bond reformation from the disulfide cleaved species. Herein, human insulin was used to investigate the performance of radical-induced reduction by the mercury lamp on larger biomolecules.

The radical-induced reduction experiments were conducted using nano electrospray ionisation (nESI) and the experimental set up is shown in Figure 3.2. Non-reduced human insulin (intact) was observed from 3+ to 6+ charge states and it is worth noting that the insulin dimer was also observed as $[2M+7H]^{7+}$ at t = 0 min when the mercury lamp was not switched on (Figure 3.21a). With mercury lamp irradiation for 5.3 min, only insulin monomer from 4+ to 6+ charge states and very low intensity of C_A and C_B were observed, suggesting the disulfide linkages were starting to be reduced and allowed more protons to attach to the insulin backbone. At t = 7.0 min, C_A and C_B were observed with slightly higher abundance, and even more C_A and C_B were observed at t = 27.0 min, showing the two inter-chain disulfide bonds were reduced effectively. However, unlike the electrochemical reduction approach, the intact insulin species were still dominant in the spectrum.

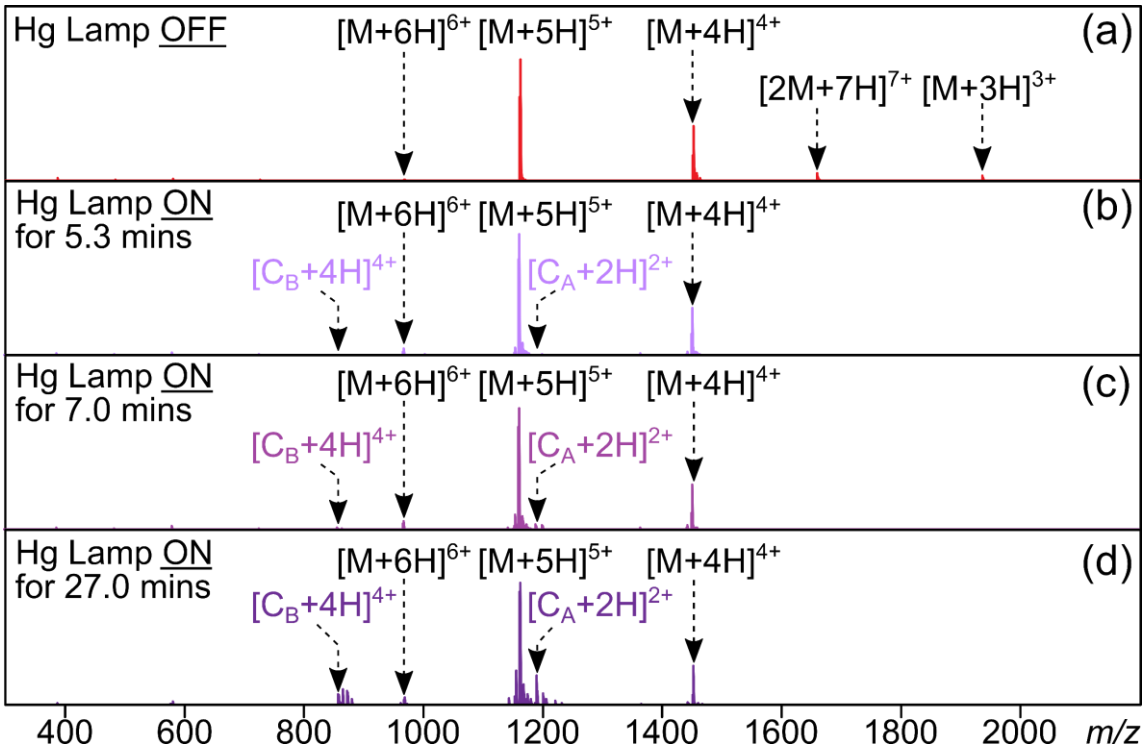


Figure 3.21 nESI-MS of human insulin. a) mercury lamp off and t = 0 min, b) mercury lamp on and t = 5.3 min, c) mercury lamp on and t = 7.0 min, and d) mercury lamp on and t = 27.0 min.

Figure 3.22 shows the nESI-MS spectrum of the insulin species when the mercury lamp been switched on for 27 min. Aside from the observation of C_A and C_B, a series of "CH_2OH" adducts were observed on different charge state species. With the aid of ultra-high resolution MS, *e.g.* FT-ICR MS, all these species were isotopically resolved, allowing accurate assignments. It is worth noting that, in some of the reduced species, S• instead of SH was observed, which could provide insight into the mechanism of the radical-induced reduction.

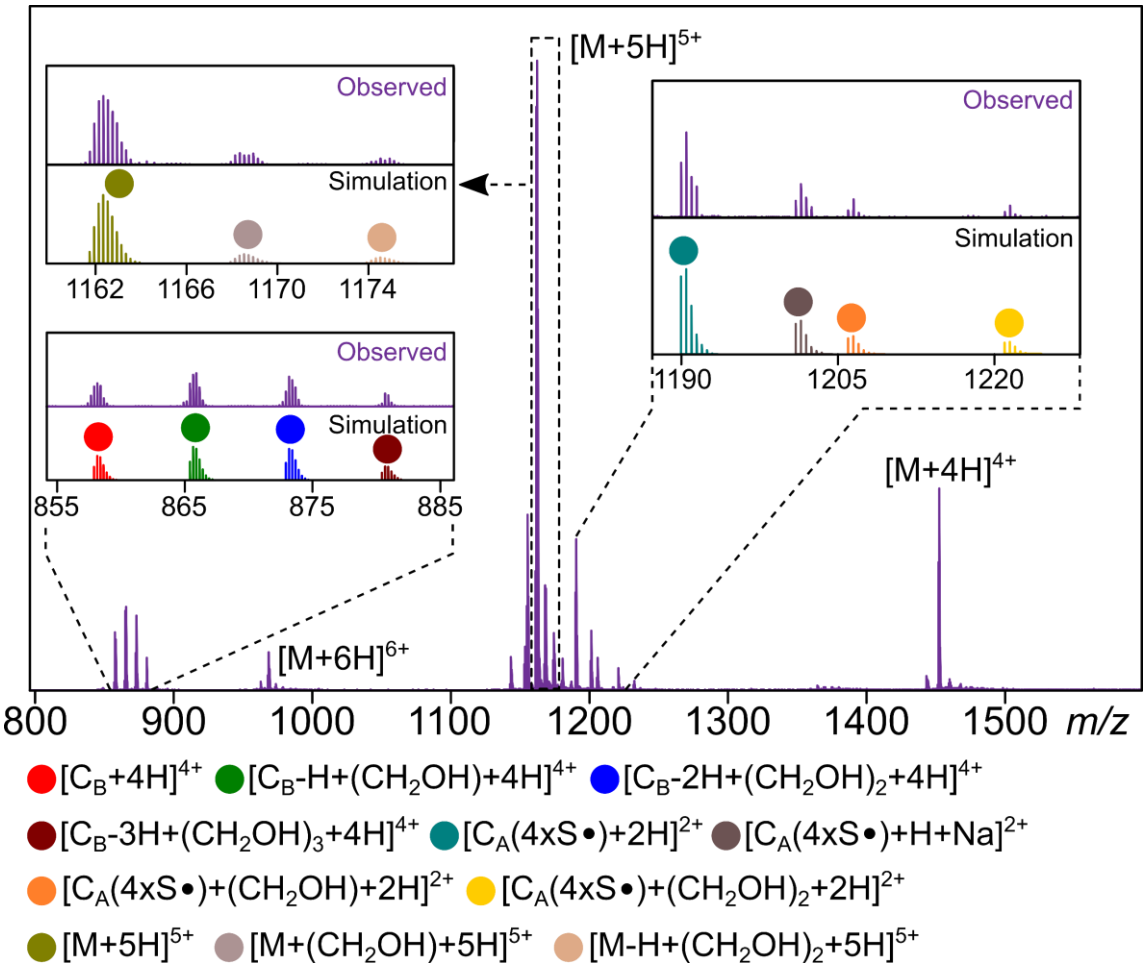


Figure 3.22 nESI-MS of human insulin (mercury lamp on for 27.0 min), a series of CH_2OH adducts were observed for different peaks (both intact and separated insulin).

MS/MS experiments were conducted to study the effect of fragmentation on non-reduced insulin. Figures 3.23 and 3.24 (Table S3.5) show the ECD MS/MS spectrum and the fragmentation map of intact insulin $[M+5H]^{5+}$ respectively. Disulfide bonds were previously shown to be preferentially cleaved with ECD MS/MS,[28] which is also observed in the fragmentation map in Figure 24. Aside from the fragments outside the "disulfide bond protected" regions, some fragments were also observed inside the regions, which the total cleavage coverage calculated to be 51%. Figure 3.25 (Table S3.6) shows the fragmentation map of the CAD MS/MS of the same species, and apart from the fragments near the termini, no fragments from inside the "protected" region were observed, as expected. The cleavage coverage was calculated to be 39%.

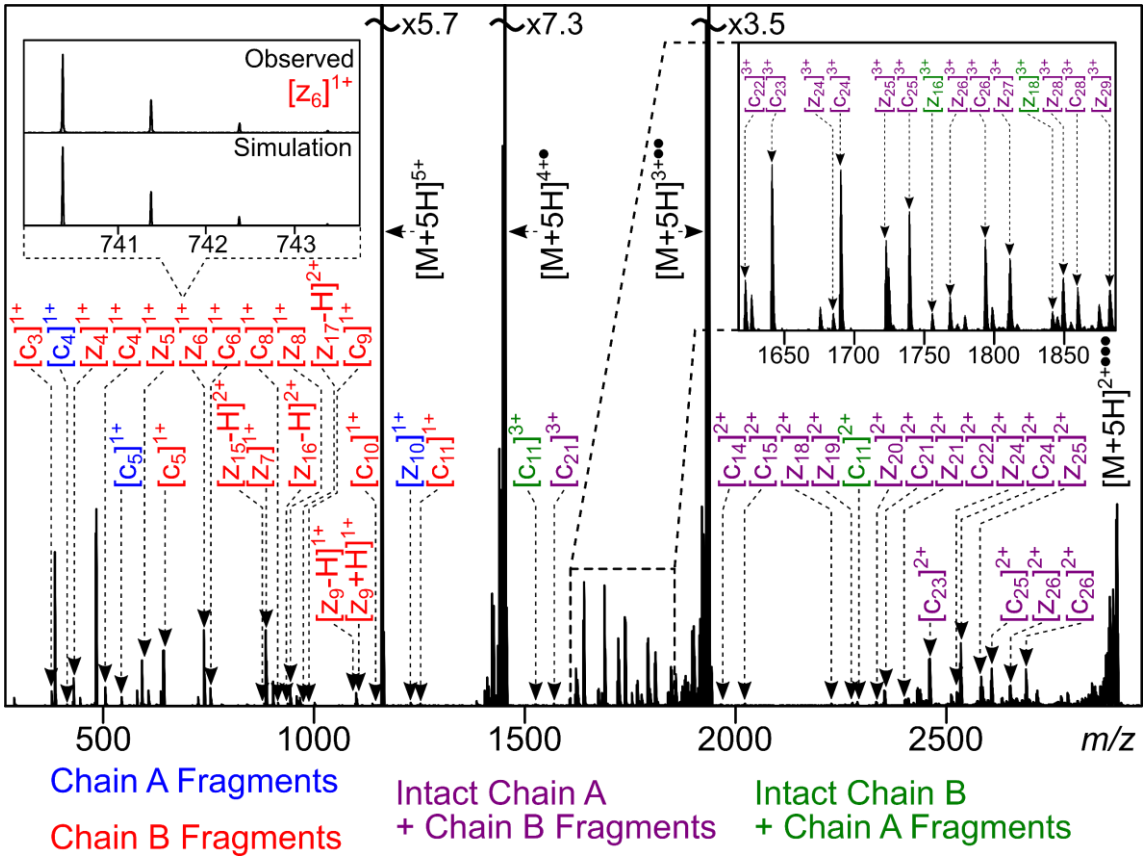


Figure 3.23 ECD MS/MS spectrum of intact insulin species $[M+5H]^{5+}$. Fragments from C_A are shown in blue, fragments from C_B are shown in red, fragments from C_B with an intact C_A are shown in purple, and fragments from C_A with an intact C_B are shown in green.

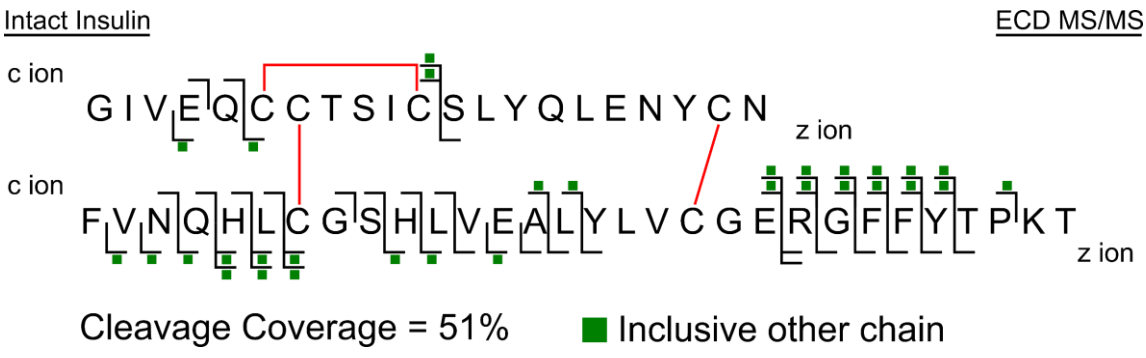


Figure 3.24 Fragmentation map of the ECD MS/MS spectrum of intact insulin species $[M+5H]^{5+}$. Fragments including the other intact chain are labelled in green. Disulfide bonds are shown in red.

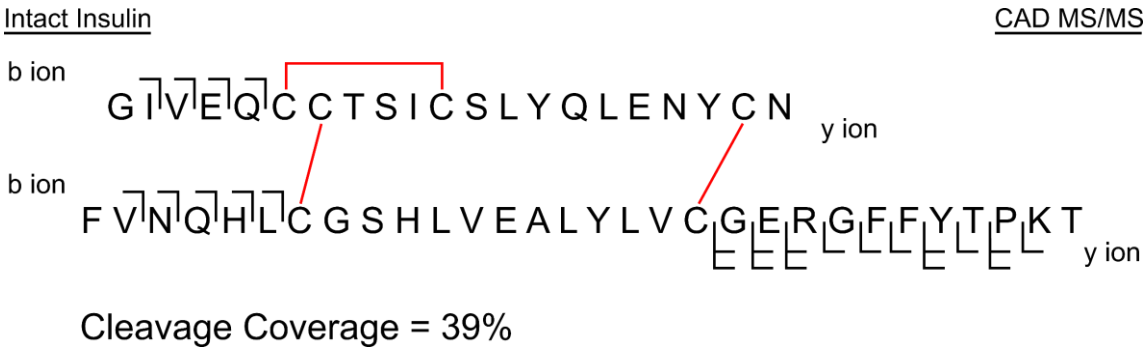


Figure 3.25 Fragmentation map of the CAD MS/MS spectrum of intact insulin species $[M+5H]^{5+}$. Disulfide bonds are shown in red.

Insulin was chemically reduced with DTT and alkylated with IAA to achieve the two separated chains (C_A and C_B). The chain B obtained was used to compare the fragmentation with the chain B produced from radical-induced reduction. Figure 3.26 (Table S3.7) shows the ECD MS/MS spectrum and the fragmentation map of the separated chain B produced from radical-induced reduction using the mercury lamp and yielded a cleavage coverage of 79%. Figures 3.27 and 3.28 (Tables S3.8 and S3.9) show the fragmentation maps of the ECD MS/MS and CAD MS/MS of the same species obtained from chemical reduction and alkylation. Interestingly, with ECD MS/MS no fragments were observed in the region between Cys7 to Cys19, even though the observation of C_B indicated C_B was no longer attached to C_A. This may suggest the two cysteine residues on C_B (originally being inter-chain disulfide bonded with C_A) have now formed a disulfide bond and protected the region, yielding a cleavage coverage of 45%. However, the fragmentation map of CAD MS/MS shows an almost complete cleavage coverage of 97%, which suggests there is no disulfide bond between the Cys7 and Cys19. This phenomenon, may be due to the non-optimum pH condition for the chemical reduction step with DTT. During the insulin sample preparation, a small volume of formic acid was needed to lower to pH of the solvent in order to dissolve all the solid. However, the reducing power of DTT is limited to pH values above 7 (basic), since only the negatively charged thiolate (S^-) is reactive and hence led to an incomplete reduction before the alkylation step with IAA.

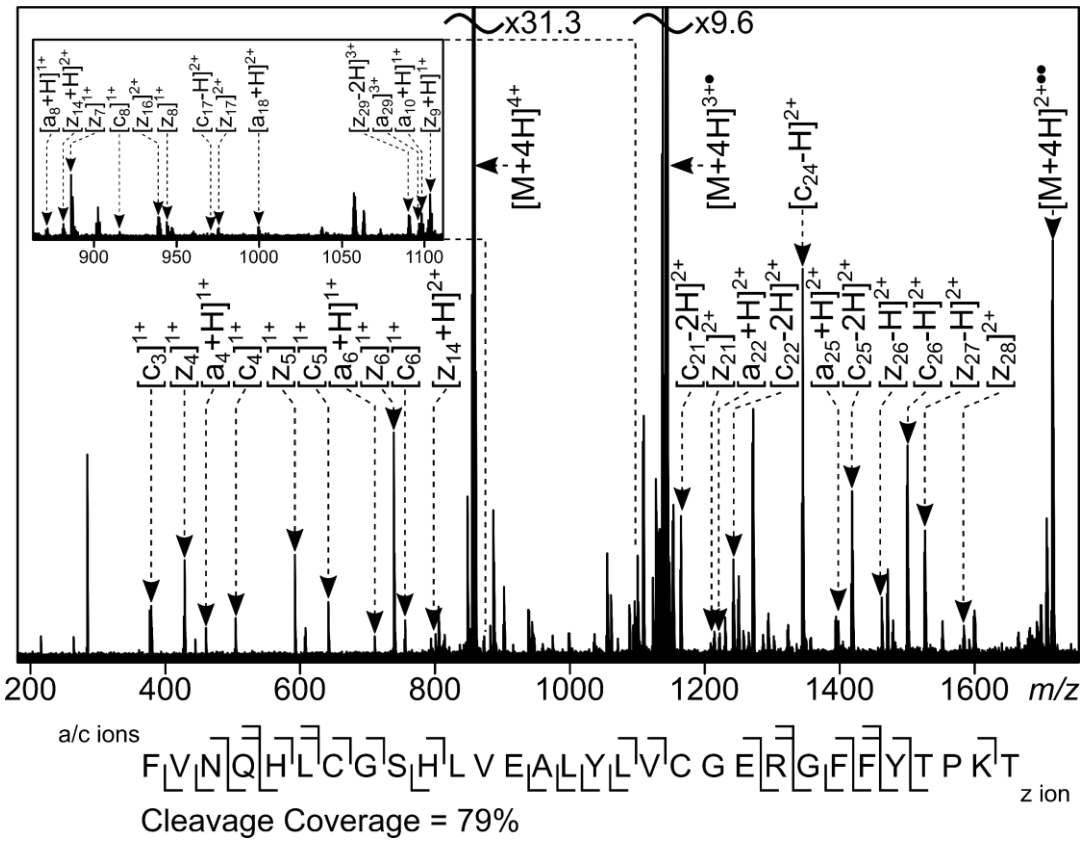


Figure 3.26 ECD MS/MS of the C_B produced from radical-induced reduction using the mercury lamp.

Figure 3.27 Fragmentation map of ECD MS/MS of the C_B produced from chemical reduction.

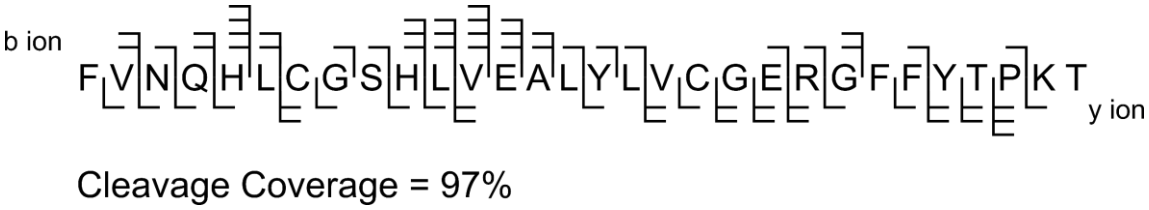


Figure 3.28 Fragmentation map of CAD MS/MS of the C_B produced from chemical reduction.

Conclusions

Disulfide bonds within protein architectures, though critical for biological function, present distinct challenges for effective MS and MS/MS characterisation of target proteins. Electrochemical reduction and radical-induced reduction of disulfide bonds provide an effective, chemical-free method for reduction which requires no further sample preparation and can provide complete disulfide bond reduction in many protein targets.

The electrochemical reduction process was monitored using human lysozyme, bovine ribonuclease A, and human insulin. The ROXY-EC reduction was shown to occur in a sequential process and could be used to reduce selectively certain disulfide bonds in stages. The reduced protein species were shown to be more effectively ionised during ESI-MS and attained much higher charge states than non-reduced species. However, the electrochemical voltage required to fully reduce a disulfide bonded protein is protein-dependent, which can be measured using a statistical method, especially with large intra-disulfide bonded proteins, where the change in mass is often minimal. Insulin also used as an example of protein with inter-chain disulfide bonds. By reducing the cysteine linkages, two separated polypeptide chains can be observed in a mass spectrum. The presence and abundance of the separated polypeptide chains could be used as an indication of percentage conversion (from intact insulin to the separated polypeptide chains). The ion abundance of the polypeptide chains was found to be the dominant peaks in the ESI-MS spectrum, overtaking the intact insulin species. In addition to monitoring the reduction process, the enhancement of MS/MS capabilities was also studied. The reduced species were shown to undergo fragmentation much more readily in subsequent MS/MS experiments. With ECD MS/MS, the reduced species was shown to have an over 7-fold increase in cleavage coverage compared to the non-reduced species, for certain proteins.

Radical-induced reduction was achieved using a mercury arc lamp, irradiating UV photons to initiate the radical-based reactions. The performance of the reduction process was shown using human insulin as a model protein. The two separated polypeptide chains were observed to increase in intensity with the irradiation time,

suggesting the reduction process occurred in the solution phase, and allowed more polypeptide chains to be observed with longer irradiation time. However, unlike EC reduction mentioned above, the intact protein species remained the dominant species in the MS spectrum. Without reducing the target protein (intact protein), the fragments observed were limited with CAD MS/MS and ECD MS/MS shown to yield better fragmentation with its inherent disulfide bond preferentiality. Chemical reduction using DTT was found to perform poorly with samples requiring low pH conditions, because DTT is only active when the sulphur is a negatively charged thiolate (S^-), which requires a basic or at least neutral pH condition to achieve. Incomplete chemical reduction was observed by analysing the CAD MS/MS and ECD MS/MS of the chain B of insulin. While using radical-induced reduction, which does not require a chemical reducing agent, like DTT, successfully reduced intact insulin to the two separated polypeptide chains, by showing a promising 79% cleavage coverage with ECD MS/MS of the chain B of insulin.

Electrochemical reduction and radical-induced reduction have been shown to be an effective tool for top-down analysis and can greatly assist in protein identification and characterisation, which is a key research area in proteomics studies, especially those containing chemically sensitive modifications, such as metallodrugs.

References

(1) Reid, G. E.; McLuckey, S. A. 'Top down' Protein Characterization *via* Tandem Mass Spectrometry. *J. Mass Spectrom.* **2002**, *37* (7), 663–675.

(2) Ntai, I.; Kim, K.; Fellers, R. T.; Skinner, O. S.; Smith, A. D.; Early, B. P.; Savaryn, J. P.; Leduc, R. D.; Thomas, P. M.; Kelleher, N. L. Applying Label-Free Quantitation to Top down Proteomics. *Anal. Chem.* **2014**, *86* (10), 4961–4968.

(3) Chen, X.; Wei, S.; Ji, Y.; Guo, X.; Yang, F. Quantitative Proteomics Using SILAC: Principles, Applications, and Developments. *Proteomics* **2015**, *15* (18), 3175–3192.

(4) Zhang, J.; Zhang, F.; Hong, C.; Giuliano, A. E.; Cui, X.; Zhou, G. Critical Protein GAPDH and Its Regulatory Mechanisms in Cancer Cells. *Cancer Biol Med* **2015**, *12* (1), 10–22.

(5) Xu, F.; Xu, Q.; Dong, X.; Guy, M.; Guner, H.; Hacker, T. A.; Ge, Y. Top-down High-Resolution Electron Capture Dissociation Mass Spectrometry for Comprehensive Characterization of Post-Translational Modifications in Rhesus Monkey Cardiac Troponin I. *Int. J. Mass Spectrom.* **2011**, *305* (2–3), 95–102.

(6) Theisen, A.; Yan, B.; Brown, J. M.; Morris, M.; Bellina, B.; Barran, P. E. Use of Ultraviolet Photodissociation Coupled with Ion Mobility Mass Spectrometry to Determine Structure and Sequence from Drift Time Selected Peptides and Proteins. *Anal. Chem.* **2016**, *88* (20), 9964–9971.

(7) Suttapitugsakul, S.; Xiao, H.; Smeekens, J.; Wu, R. Evaluation and Optimization of Reduction and Alkylation Methods to Maximize Peptide Identification with MS-Based Proteomics. *Mol. Biosyst.* **2017**, *13* (12), 2574–2582.

(8) Chen, B.; Brown, K. A.; Lin, Z.; Ge, Y. Top-Down Proteomics: Ready for Prime Time? *Anal. Chem.* **2018**, *90* (1), 110–127.

(9) Thornton, J. M. Disulfide Bridges in Globular Proteins. *J. Mol. Biol.* **1981**, *151*, 261–287.

(10) Xia, Y.; Cooks, R. G. Plasma Induced Oxidative Cleavage of Disulfide Bonds in Polypeptides during Nanoelectrospray Ionization. *Anal. Chem.* **2010**, *82* (7), 2856–2864.

(11) Mysling, S.; Salbo, R.; Ploug, M.; Jørgensen, T. J. D. Electrochemical Reduction of Disulfide-Containing Proteins for Hydrogen/Deuterium Exchange Monitored by Mass Spectrometry. *Anal. Chem.* **2014**, *86* (1), 340–345.

(12) Switzar, L.; Nicolardi, S.; Rutten, J. W.; Oberstein, S. A. J. L.; Aartsma-Rus, A.; Van Der Burgt, Y. E. M. In-Depth Characterization of Protein Disulfide Bonds by Online Liquid Chromatography-Electrochemistry-Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2016**, *27* (1), 50–58.

(13) Tan, L.; Xia, Y. Gas-Phase Reactivity of Peptide Thiyl (RS•), Perthiyl (RSS•), and Sulfinyl (RSO•) Radical Ions Formed from Atmospheric Pressure Ion/Radical Reactions. *J. Am. Soc. Mass Spectrom.* **2013**, *24* (4), 534–542.

(14) Stinson, C. A.; Xia, Y. Reactions of Hydroxyalkyl Radicals with Cysteinyl Peptides in a NanoESI Plume. *J Am Soc Mass Spectrom* **2014**, *25*, 1192–1201.

(15) Wei, J.; Antzutkin, O. N.; Filippov, A. V.; Iuga, D.; Lam, P. Y.; Barrow, M. P.; Dupree, R.; Brown, S. P.; O'Connor, P. B. Amyloid Hydrogen Bonding Polymorphism Evaluated by 15 N{ 17 O}REAPDOR Solid-State NMR and Ultra-High Resolution Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Biochemistry* **2016**, *55* (14), 2065–2068.

(16) Qi, Y.; Barrow, M. P.; Li, H.; Meier, J. E.; Orden, S. L. Van; Thompson, C. J.; Connor, P. B. O. Absorption-Mode: The Next Generation of Fourier Transform Mass Spectra. *Anal. Chem.* **2012**, *84*, 2923–2929.

(17) Jennings, K. R. Colllision-Induced Decompositions of Aromatic Ions. *Int. J. Mass Spectrom. Ion Phys.* **1968**, *1* (3), 227–235.

(18) Zubarev, R.; Kelleher, N. L.; McLafferty, F. W. Electron Capture Dissociation of Multiply Charged Protein Cations. A Nonergodic Process. *J. Am. Chem. Soc* **1998**, *120* (16), 3265–3266.

(19) Nguyen, V. H.; Afonso, C.; Tabet, J. C. Comparison of Collision-Induced Dissociation and Electron-Induced Dissociation of Singly Charged Mononucleotides. *Int. J. Mass Spectrom.* **2012**, *316–318*, 140–146.

(20) Wootton, C. A.; Sanchez-Cano, C.; Liu, H.-K.; Barrow, M. P.; Sadler, P. J.; O'Connor, P. B. Binding of an Organo-Osmium(II) Anticancer Complex to Guanine and Cytosine on DNA Revealed by Electron-Based Dissociations in High Resolution Top-Down FT-ICR Mass Spectrometry. *Dalton Trans.* **2015**, *44* (8), 3624–3632.

(21) Little, D. P.; Speir, J. P.; Senko, M. W.; O'Connor, P. B.; McLafferty, F. W. Infrared Multiphoton Dissociation of Large Multiply Charged Ions for Biomolecule Sequencing. *Anal. Chem.* **1994**, *66* (18), 2809–2815.

(22) Shaw, J. B.; Li, W.; Holden, D. D.; Zhang, Y.; Griep-Raming, J.; Fellers, R. T.; Early, B. P.; Thomas, P. M.; Kelleher, N. L.; Brodbelt, J. S. Complete Protein Characterization Using Top-down Mass Spectrometry and Ultraviolet Photodissociation. *J. Am. Chem. Soc.* **2013**, *135* (34), 12646–12651.

(23) Laskin, J.; Futrell, J. H. Activation of Large Ions in FT-ICR Mass Spectrometry. *Mass Spectrom. Rev.* **2005**, *24* (2), 135–167.

(24) Caravatti, P.; Allemann, M. The 'Infinity Cell': A New Trapped-ion Cell with Radiofrequency Covered Trapping Electrodes for Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Org. Mass Spectrom.* **1991**, *26* (5), 514–518.

(25) Zhang, L.; Rempel, D.; Pramanik, B. N.; Gross, M. L. Accurate Mass Measurements by Fourier Transform Mass Spectrometry. *Mass Spectrom. Rev.* **2005**, *24*, 286–309.

(26) Kraj, A.; Brouwer, H. J.; Reinhoud, N.; Chervet, J. P. A Novel Electrochemical Method for Efficient Reduction of Disulfide Bonds in Peptides and Proteins Prior to MS Detection. *Anal. Bioanal. Chem.* **2013**, *405* (29), 9311–9320.

(27) Trabjerg, E.; Jakobsen, R. U.; Mysling, S.; Christensen, S.; Jørgensen, T. J. D.;

Rand, K. D. Conformational Analysis of Large and Highly Disulfide-Stabilized Proteins by Integrating Online Electrochemical Reduction into an Optimized H/D Exchange Mass Spectrometry Workflow. *Anal. Chem.* **2015**, *87* (17), 8880–8888.

(28) Zhurov, K. O.; Fornelli, L.; Wodrich, M. D.; Laskay, A.; Tsybin, Y. O. Principles of Electron Capture and Transfer Dissociation Mass Spectrometry Applied to Peptide and Protein Structure Analysis. *Chem. Soc. Rev. Chem. Soc. Rev* **2013**, *42* (42), 5014–5030.

Supplementary Information

Table S3.1 Assignments for Figures 3.11 and 3.12: CAD MS/MS of the non-reduced hLZ species.

Assignment	Charge State	Elemental Composition	Theoretical Mass (m/z)	Observed Mass (m/z)	Mass Error (ppm)
b19	3+	$C_{95}H_{165}N_{30}O_{26}S_2$	735.397877	735.39787	-0.01
b6-H_2O	1+	$C_{34}H_{53}N_{10}O_7S_1$	745.381392	745.38176	0.49
MH-H_2O	12+	$C_{633}H_{1002}N_{200}O_{185}S_{10}$	1223.769068	1223.76911	0.03
MH-H_2O	11+	$C_{633}H_{1001}N_{200}O_{185}S_{10}$	1334.929231	1334.92932	0.07
MH-NH_3	11+	$C_{633}H_{1000}N_{199}O_{186}S_{10}$	1335.018687	1335.01808	-0.45
b123	10+	$C_{602}H_{954}N_{192}O_{176}S_9$	1397.690345	1397.69026	-0.06
y125	10+	$C_{602}H_{953}N_{191}O_{179}S_{10}$	1404.184936	1404.18553	0.42
y128-H_2O	10+	$C_{622}H_{979}N_{197}O_{183}S_{10}$	1445.605092	1445.60559	0.34
b128-H_2O	10+	$C_{626}H_{986}N_{198}O_{182}S_{10}$	1450.911385	1450.91165	0.18
y124-H_2O	9+	$C_{599}H_{945}N_{190}O_{177}S_9$	1546.646926	1546.64608	-0.55
y124	9+	$C_{599}H_{947}N_{190}O_{178}S_9$	1548.648100	1548.64868	0.37
y125	9+	$C_{602}H_{952}N_{191}O_{179}S_{10}$	1560.093565	1560.09325	-0.20
b124-NH_3	9+	$C_{611}H_{959}N_{192}O_{178}S_9$	1569.103661	1569.10395	0.18
y126-H_2O	9+	$C_{608}H_{962}N_{195}O_{179}S_{10}$	1575.436959	1575.43681	-0.09
y126	9+	$C_{608}H_{964}N_{195}O_{180}S_{10}$	1577.438133	1577.43838	0.16
b126	9+	$C_{621}H_{979}N_{196}O_{181}S_9$	1596.231832	1596.23143	-0.25
b127-H_2O	9+	$C_{623}H_{980}N_{197}O_{181}S_9$	1600.566377	1600.56693	0.35
b128-H_2O	9+	$C_{626}H_{985}N_{198}O_{182}S_{10}$	1612.011842	1612.01158	-0.16
y129	9+	$C_{627}H_{989}N_{198}O_{185}S_{10}$	1619.124736	1619.12435	-0.24
b14	1+	$C_{76}H_{132}N_{25}O_{19}S_1$	1730.984657	1730.98401	-0.37
y124-H_2O	8+	$C_{599}H_{944}N_{190}O_{177}S_9$	1739.851883	1739.85178	-0.06
b124	8+	$C_{611}H_{961}N_{193}O_{178}S_9$	1767.244028	1767.24424	0.12
b125-H_2O	8+	$C_{616}H_{968}N_{194}O_{178}S_9$	1777.376259	1777.37626	0.00
b125	8+	$C_{616}H_{970}N_{194}O_{179}S_9$	1779.627580	1779.62788	0.17
y16-NH_3	1+	$C_{75}H_{123}N_{28}O_{24}S_2$	1863.870100	1863.87162	0.82
b124	7+	$C_{611}H_{960}N_{193}O_{178}S_9$	2019.563564	2019.56373	0.08
				Average	0.05 ppm
				Absolute Average	0.24 ppm
				Standard Deviation	0.31 ppm

Table S3.2 Assignments for Figures 3.13 and 3.14: CAD MS/MS of the EC-reduced hLZ species.

Assignment	Charge State	Elemental Composition	Theoretical Mass (m/z)	Observed Mass (m/z)	Mass Error (ppm)
b5	1+	$C_{31}H_{50}N_9O_7$	660.382772	660.38294	0.25
b26	4+	$C_{130}H_{222}N_{40}O_{35}S_2$	741.906020	741.90614	0.16
b6	1+	$C_{34}H_{55}N_{10}O_8S_1$	763.391957	763.39178	-0.23
b27	4+	$C_{134}H_{228}N_{42}O_{37}S_2$	770.416752	770.41697	0.28
y41	6+	$C_{198}H_{327}N_{71}O_{55}S_3$	779.235046	779.23454	-0.65
b13	2+	$C_{70}H_{121}N_{21}O_{18}S_1$	787.945411	787.94533	-0.10
b28	4+	$C_{145}H_{238}N_{44}O_{38}S_2$	816.936580	816.93634	-0.29
y28	4+	$C_{142}H_{227}N_{51}O_{38}S_2$	829.670441	829.66993	-0.62
y44	6+	$C_{212}H_{349}N_{75}O_{61}S_3$	836.260701	836.26126	0.67
MH-H_2O	17+	$C_{633}H_{1007}N_{200}O_{185}S_{10}$	864.133247	864.13358	0.39
MH	17+	$C_{633}H_{1009}N_{200}O_{186}S_{10}$	865.192692	865.19268	-0.01
b7-H_2O	1+	$C_{39}H_{60}N_{11}O_{10}S_1$	874.423986	874.42321	-0.89
b7	1+	$C_{39}H_{62}N_{11}O_{11}S_1$	892.434551	892.43437	-0.20
b55	7+	$C_{273}H_{429}N_{82}O_{82}S_4$	899.439495	899.43864	-0.95
MH	16+	$C_{633}H_{1008}N_{200}O_{186}S_{10}$	919.204280	919.20448	0.22
y31	4+	$C_{157}H_{253}N_{57}O_{43}S_2$	922.219558	922.21987	0.34
y101	12+	$C_{483}H_{761}N_{155}O_{147}S_7$	942.373476	942.37364	0.17
MH	15+	$C_{633}H_{1007}N_{200}O_{186}S_{10}$	980.417414	980.41679	-0.64
y101	11+	$C_{483}H_{760}N_{155}O_{147}S_7$	1027.952221	1027.95265	0.42
MH	14+	$C_{633}H_{1006}N_{200}O_{186}S_{10}$	1050.375281	1050.37536	0.08
y18	2+	$C_{85}H_{146}N_{35}O_{27}S_2$	1076.527895	1076.52683	-0.99
y10	1+	$C_{47}H_{78}N_{15}O_{14}S_1$	1108.556792	1108.55583	-0.87
MH	13+	$C_{633}H_{1005}N_{200}O_{186}S_{10}$	1131.091897	1131.09121	-0.61
b10-H_2O	1+	$C_{54}H_{88}N_{17}O_{13}S_1$	1214.646275	1214.64706	0.65
MH	12+	$C_{633}H_{1004}N_{200}O_{186}S_{10}$	1225.269948	1225.27064	0.56
MH	11+	$C_{633}H_{1003}N_{200}O_{186}S_{10}$	1336.566555	1336.56689	0.25
b12-H_2O	1+	$C_{64}H_{106}N_{19}O_{16}S_1$	1428.778018	1428.77694	-0.75
MH-H_2O	10+	$C_{633}H_{1000}N_{200}O_{185}S_{10}$	1468.321426	1468.32170	0.19
MH	10+	$C_{633}H_{1002}N_{200}O_{186}S_{10}$	1470.122483	1470.12270	0.15
y13	1+	$C_{61}H_{101}N_{22}O_{20}S_1$	1493.727775	1493.72629	-0.99
y14-NH_3	1+	$C_{66}H_{106}N_{23}O_{22}S_1$	1604.759804	1604.75908	-0.45
y14	1+	$C_{66}H_{109}N_{24}O_{22}S_1$	1621.786353	1621.78488	-0.91
MH-H_2O	9+	$C_{633}H_{999}N_{200}O_{185}S_{10}$	1631.356332	1631.35770	0.84
MH	9+	$C_{633}H_{1001}N_{200}O_{186}S_{10}$	1633.357506	1633.35746	-0.03
y43	3+	$C_{208}H_{341}N_{74}O_{58}S_3$	1633.171812	1633.17136	-0.28
y29-H_2O	2+	$C_{146}H_{228}N_{52}O_{40}S_2$	1706.841796	1706.84140	-0.23
y29	2+	$C_{146}H_{230}N_{52}O_{41}S_2$	1715.847078	1715.84738	0.18

MH	8+	$C_{633}H_{1000}N_{200}O_{186}S_{10}$	1837.401284	1837.40166	0.20
b35-NH_3	2+	$C_{184}H_{292}N_{52}O_{47}S_4$	2055.046470	2055.04633	-0.07
				Average	-0.12 ppm
				Absolute Average	0.43 ppm
				Standard Deviation	0.52 ppm

Table S3.3 Assignments for Figures 3.15 and 3.16: ECD MS/MS of the non-reduced hLZ species.

Assignment	Charge State	Elemental Composition	Theoretical Mass (m/z)	Observed Mass (m/z)	Mass Error (ppm)
c3	1+	$C_{20}H_{34}N_5O_3$	392.265616	392.26557	-0.12
c4	1+	$C_{25}H_{41}N_6O_6$	521.308210	521.30850	0.56
[c+1]78	8+	$C_{386}H_{600}N_{117}O_{114}S_6$	1111.287873	1111.28695	-0.83
c78	7+	$C_{386}H_{598}N_{117}O_{114}S_6$	1269.755414	1269.75569	0.22
MH	11+	$C_{633}H_{1003}N_{200}O_{186}S_{10}$	1336.566555	1336.56586	-0.52
z124	9+	$C_{599}H_{945}N_{189}O_{178}S_9$	1546.868242	1546.86921	0.63
				Average	-0.01 ppm
				Absolute Average	0.48 ppm
				Standard Deviation	0.59 ppm

Table S3.4 Assignments for Figures 3.17 and 3.18: ECD MS/MS of the EC-reduced hLZ species.

Assignment	Charge State	Elemental Composition	Theoretical Mass (m/z)	Observed Mass (m/z)	Mass Error (ppm)
c3	1+	$C_{20}H_{34}N_5O_3$	392.265616	392.26565	0.09
c17	4+	$C_{89}H_{161}N_{29}O_{22}S_2$	513.044762	513.04451	-0.49
c4	1+	$C_{25}H_{41}N_6O_6$	521.308210	521.30835	0.27
c14	3+	$C_{76}H_{137}N_{26}O_{19}S_1$	583.341919	583.34164	-0.48
c20	4+	$C_{104}H_{178}N_{32}O_{28}S_2$	596.822696	596.82232	-0.63
c10	2+	$C_{54}H_{94}N_{18}O_{14}S_1$	625.345333	625.34535	0.03
c21	4+	$C_{110}H_{190}N_{36}O_{29}S_2$	635.847974	635.84808	0.17
c16	3+	$C_{84}H_{151}N_{28}O_{21}S_1$	640.043762	640.04384	0.12
c22	4+	$C_{112}H_{193}N_{37}O_{30}S_2$	650.103340	650.10357	0.35
c11	2+	$C_{58}H_{101}N_{19}O_{16}S_1$	675.869172	675.86941	0.35
c5	1+	$C_{31}H_{53}N_{10}O_7$	677.409321	677.40901	-0.46
c23	4+	$C_{118}H_{204}N_{38}O_{31}S_2$	678.374356	678.37413	-0.33
c17	3+	$C_{89}H_{160}N_{29}O_{22}S_2$	683.723924	683.72390	-0.04
c25	4+	$C_{127}H_{220}N_{40}O_{34}S_2$	728.403379	728.40344	0.08
c12	2+	$C_{64}H_{112}N_{20}O_{17}S_1$	732.411204	732.41152	0.43
c26	4+	$C_{130}H_{225}N_{41}O_{35}S_2$	746.162657	746.16236	-0.40
c6	1+	$C_{34}H_{58}N_{11}O_8S_1$	780.418506	780.41818	-0.42
c20	3+	$C_{104}H_{177}N_{32}O_{28}S_2$	795.427836	795.42772	-0.15
c13	2+	$C_{70}H_{124}N_{22}O_{18}S_1$	796.458686	796.45876	0.09
c28	4+	$C_{145}H_{241}N_{45}O_{38}S_2$	821.193218	821.19310	-0.14
c35	5+	$C_{184}H_{301}N_{54}O_{47}S_4$	829.433573	829.43348	-0.11
c43	6+	$C_{219}H_{356}N_{66}O_{60}S_4$	833.094746	833.09491	0.20
c36	5+	$C_{187}H_{306}N_{55}O_{49}S_4$	846.839979	846.83962	-0.42
c29	4+	$C_{150}H_{250}N_{46}O_{39}S_3$	853.953339	853.95331	-0.03
c14	2+	$C_{76}H_{136}N_{26}O_{19}S_1$	874.509241	874.50926	0.02
c45	6+	$C_{232}H_{371}N_{69}O_{64}S_4$	879.279122	879.27931	0.21
c23	3+	$C_{118}H_{203}N_{38}O_{31}S_2$	904.163382	904.16316	-0.25
c7	1+	$C_{39}H_{65}N_{12}O_{11}S_1$	909.461100	909.46152	0.46
c32	4+	$C_{162}H_{271}N_{49}O_{42}S_4$	925.735930	925.73623	0.32
[c+1]24	3+	$C_{121}H_{209}N_{39}O_{33}S_2$	933.509999	933.50932	-0.73
c41	5+	$C_{212}H_{343}N_{64}O_{57}S_4$	965.095281	965.09562	0.35
c25	3+	$C_{127}H_{219}N_{40}O_{34}S_2$	970.868747	970.86807	-0.70
MH	15+	$C_{633}H_{1007}N_{200}O_{186}S_{10}$	980.417414	980.41798	0.58
c26	3+	$C_{130}H_{224}N_{41}O_{35}S_2$	994.547785	994.54743	-0.36
c43	5+	$C_{219}H_{355}N_{66}O_{60}S_4$	999.512240	999.51237	0.13
c17	2+	$C_{89}H_{159}N_{29}O_{22}S_2$	1025.082247	1025.08234	0.09
c35	4+	$C_{184}H_{300}N_{54}O_{47}S_4$	1036.540147	1036.54078	0.61

c54	6+	$C_{271}H_{428}N_{82}O_{81}S_4$	1042.512379	1042.51257	0.18
MH	14+	$C_{633}H_{1006}N_{200}O_{186}S_{10}$	1050.375281	1050.37522	-0.06
[z+2]36	4+	$C_{180}H_{296}N_{65}O_{48}S_3$	1058.046485	1058.04646	-0.02
c9	1+	$C_{48}H_{81}N_{14}O_{13}S_1$	1093.582278	1093.58211	-0.15
c19	2+	$C_{95}H_{167}N_{31}O_{26}S_2$	1111.106451	1111.10668	0.21
MH	13+	$C_{633}H_{1005}N_{200}O_{186}S_{10}$	1131.095897	1131.09548	-0.37
c20	2+	$C_{104}H_{176}N_{32}O_{28}S_2$	1192.638116	1192.63829	0.15
MH	12+	$C_{633}H_{1004}N_{200}O_{186}S_{10}$	1225.269948	1225.26964	-0.25
[c+2]78	7+	$C_{386}H_{600}N_{117}O_{114}S_6$	1270.043361	1270.04391	0.43
[c-1]124	11+	$C_{611}H_{966}N_{194}O_{178}S_9$	1286.999340	1286.99923	-0.09
MH	11+	$C_{633}H_{1003}N_{200}O_{186}S_{10}$	1336.566555	1336.56669	0.10
MH	10+	$C_{633}H_{1002}N_{200}O_{186}S_{10}$	1470.122483	1470.12297	0.33
z124	9+	$C_{599}H_{945}N_{189}O_{178}S_9$	1546.868242	1546.86836	0.08
MH	9+	$C_{633}H_{1001}N_{200}O_{186}S_{10}$	1633.357506	1633.35720	-0.19
[z+3]126	8+	$C_{608}H_{964}N_{194}O_{180}S_{10}$	1772.867580	1772.86700	-0.33
MH	8+	$C_{633}H_{1000}N_{200}O_{186}S_{10}$	1837.401284	1837.40228	0.54
MH	7+	$C_{633}H_{999}N_{200}O_{186}S_{10}$	2099.743285	2099.74403	0.35
MH	6+	$C_{633}H_{998}N_{200}O_{186}S_{10}$	2449.532620	2449.53445	0.75
				Average	0.01 ppm
				Absolute Average	0.28 ppm
				Standard Deviation	0.35 ppm

Table S3.5 Assignments for Figures 3.23 and 3.24: ECD MS/MS of the intact insulin species $[M+5H]^{5+}$.

Assignment	Charge State	Elemental Composition	Theoretical Mass (m/z)	Observed Mass (m/z)	Mass Error (ppm)
c3(CB)	1+	$C_{18}H_{28}N_5O_4$	378.213581	378.21351	-0.19
c4(CA)	1+	$C_{18}H_{34}N_5O_6$	416.250361	416.25026	-0.24
z4(CB)	1+	$C_{19}H_{34}N_4O_7$	430.242202	430.24205	-0.35
c4(CB)	1+	$C_{23}H_{36}N_7O_6$	506.272159	506.27201	-0.29
c5(CA)	1+	$C_{23}H_{42}N_7O_8$	544.308939	544.30882	-0.22
z5(CB)	1+	$C_{28}H_{43}N_5O_9$	593.305531	593.30544	-0.15
c5(CB)	1+	$C_{29}H_{43}N_{10}O_7$	643.331071	643.33104	-0.05
z6(CB)	1+	$C_{37}H_{52}N_6O_{10}$	740.373945	740.37401	0.09
c6(CB)	1+	$C_{35}H_{54}N_{11}O_8$	756.415135	756.41532	0.24
z15-H(CB)	2+	$C_{84}H_{120}N_{18}O_{22}S_1$	882.426717	882.42636	-0.40
z7(CB)	1+	$C_{46}H_{61}N_7O_{11}$	887.442359	887.44260	0.27
c8(CB)	1+	$C_{40}H_{62}N_{13}O_{10}S_1$	916.445784	916.44578	0.00
z16-H(CB)	2+	$C_{90}H_{131}N_{19}O_{23}S_1$	938.968750	938.96879	0.04
z8(CB)	1+	$C_{48}H_{64}N_8O_{12}$	944.463823	944.46368	-0.15
z17-H(CB)	2+	$C_{93}H_{136}N_{20}O_{24}S_1$	974.487307	974.48677	-0.55
c9(CB)	1+	$C_{43}H_{67}N_{14}O_{12}S_1$	1003.477813	1003.47808	0.27
z9-H(CB)	1+	$C_{54}H_{75}N_{12}O_{13}$	1099.557109	1099.55607	-0.94
z9+H(CB)	1+	$C_{54}H_{77}N_{12}O_{13}$	1101.572759	1101.57289	0.12
c10(CB)	1+	$C_{49}H_{74}N_{17}O_{13}S_1$	1140.536725	1140.53724	0.45
z10(CA)	1+	$C_{54}H_{78}N_{12}O_{19}S_1$	1230.522145	1230.52251	0.30
c11(CB)	1+	$C_{55}H_{85}N_{18}O_{14}S_1$	1253.620789	1253.62187	0.86
c11+CB	3+	$C_{203}H_{311}N_{53}O_{58}S_5$	1526.386755	1526.38628	-0.31
c21+CA	3+	$C_{203}H_{312}N_{53}O_{64}S_6$	1569.369883	1569.37058	0.44
c22+CA	3+	$C_{209}H_{324}N_{57}O_{65}S_6$	1621.403586	1621.40442	0.51
c23+CA	3+	$C_{211}H_{327}N_{58}O_{66}S_6$	1640.410741	1640.41209	0.82
z24+CA	3+	$C_{222}H_{334}N_{54}O_{69}S_6$	1684.086482	1684.08725	0.46
c24+CA	3+	$C_{220}H_{336}N_{59}O_{67}S_6$	1689.433545	1689.43309	-0.27
z25+CA	3+	$C_{228}H_{345}N_{55}O_{70}S_6$	1721.781170	1721.78259	0.82
c25+CA	3+	$C_{229}H_{345}N_{60}O_{68}S_6$	1738.456350	1738.45711	0.44
z16+CB	3+	$C_{234}H_{346}N_{58}O_{69}S_6$	1754.788548	1754.78991	0.78
z26+CA	3+	$C_{234}H_{352}N_{58}O_{71}S_6$	1767.467474	1767.46820	0.41
c26+CA	3+	$C_{238}H_{354}N_{61}O_{70}S_6$	1792.810793	1792.80986	-0.52
z27+CA	3+	$C_{239}H_{360}N_{60}O_{73}S_6$	1810.153666	1810.15287	-0.44
z18+CB	3+	$C_{244}H_{361}N_{61}O_{74}S_6$	1840.488938	1840.49057	0.89
z28+CA	3+	$C_{243}H_{366}N_{62}O_{75}S_6$	1848.167976	1848.16751	-0.25
c28+CA	3+	$C_{247}H_{368}N_{63}O_{73}S_6$	1858.844274	1858.84415	-0.07
z29+CA	3+	$C_{248}H_{375}N_{63}O_{76}S_6$	1881.190780	1881.19254	0.94

c14+CA	2+	$C_{167}H_{257}N_{46}O_{54}S_5$	1965.368542	1965.36759	-0.48
c15+CA	2+	$C_{173}H_{268}N_{47}O_{55}S_5$	2021.910574	2021.90948	-0.54
z18+CA	2+	$C_{197}H_{294}N_{46}O_{62}S_5$	2227.992964	2227.99250	-0.21
z19+CA	2+	$C_{202}H_{303}N_{47}O_{63}S_5$	2277.527171	2277.52659	-0.26
c11+CB	2+	$C_{203}H_{310}N_{53}O_{58}S_5$	2289.076494	2289.07793	0.63
z20+CA	2+	$C_{208}H_{314}N_{48}O_{64}S_5$	2334.069203	2334.06976	0.24
c21+CA	2+	$C_{203}H_{311}N_{53}O_{64}S_6$	2353.551186	2353.55291	0.73
z21+CA	2+	$C_{214}H_{321}N_{51}O_{65}S_5$	2402.598659	2402.59685	-0.75
c22+CA	2+	$C_{209}H_{323}N_{57}O_{65}S_6$	2431.601741	2431.60240	0.27
c23+CA	2+	$C_{211}H_{326}N_{58}O_{66}S_6$	2460.112473	2460.11346	0.40
z24+CA	2+	$C_{222}H_{333}N_{54}O_{69}S_6$	2525.626085	2525.62843	0.93
c24+CA	2+	$C_{220}H_{335}N_{59}O_{67}S_6$	2533.646680	2533.64533	-0.53
z25+CA	2+	$C_{228}H_{344}N_{55}O_{70}S_6$	2582.168117	2582.16989	0.69
c25+CA	2+	$C_{229}H_{344}N_{60}O_{68}S_6$	2607.180887	2607.17911	-0.68
z26+CA	2+	$C_{234}H_{351}N_{58}O_{71}S_6$	2650.697573	2650.70018	0.98
c26+CA	2+	$C_{238}H_{353}N_{61}O_{70}S_6$	2688.712551	2688.71455	0.74
				Average	0.11 ppm
				Absolute Average	0.45 ppm
				Standard Deviation	0.51 ppm

Table S3.6 Assignments for Figure 3.25: CAD MS/MS of the intact insulin species $[M+5H]^{5+}$.

Assignment	Charge State	Elemental Composition	Theoretical Mass (m/z)	Observed Mass (m/z)	Mass Error (ppm)
b2(CA)	1+	$C_8H_{15}N_2O_2$	171.112804	171.11280	-0.02
b2(CB)	1+	$C_{14}H_{19}N_2O_2$	247.144104	247.14413	0.11
y2(CB)	1+	$C_{10}H_{22}N_3O_4$	248.160483	248.16049	0.03
b3(CA)	1+	$C_{13}H_{24}N_3O_3$	270.181218	270.18121	-0.03
y3-H_2O(CB)	1+	$C_{15}H_{27}N_4O_4$	327.202682	327.20263	-0.16
y3(CB)	1+	$C_{15}H_{29}N_4O_5$	345.213247	345.21343	0.53
b3(CB)	1+	$C_{18}H_{25}N_4O_4$	361.187032	361.18728	0.69
b4(CA)	1+	$C_{18}H_{31}N_4O_6$	399.223812	399.22370	-0.28
y4(CB)	1+	$C_{19}H_{36}N_5O_7$	446.260926	446.26111	0.41
b4(CB)	1+	$C_{23}H_{33}N_6O_6$	489.245610	489.24538	-0.47
b5(CA)	1+	$C_{23}H_{39}N_6O_8$	527.282390	527.28196	-0.82
y9-H_2O(CB)	2+	$C_{54}H_{77}N_{13}O_{12}$	549.790185	549.79058	0.72
y9(CB)	2+	$C_{54}H_{79}N_{13}O_{13}$	558.795467	558.79543	-0.07
y5-H_2O(CB)	1+	$C_{28}H_{43}N_6O_8$	591.313690	591.31326	-0.73
b5-CO(CB)	1+	$C_{28}H_{40}N_9O_6$	598.309607	598.30962	0.02
y5(CB)	1+	$C_{28}H_{45}N_6O_9$	609.324255	609.32458	0.53
y10-H_2O(CB)	2+	$C_{59}H_{84}N_{14}O_{15}$	614.311482	614.31130	-0.30
y10(CB)	2+	$C_{59}H_{86}N_{14}O_{16}$	623.316764	623.31726	0.80
y11-H_2O(CB)	2+	$C_{61}H_{87}N_{15}O_{16}$	642.822214	642.82226	0.07
y11(CB)	2+	$C_{61}H_{89}N_{15}O_{17}$	651.827496	651.82777	0.42
b6(CB)	1+	$C_{35}H_{51}N_{10}O_8$	739.388586	739.38898	0.53
y6(CB)	1+	$C_{37}H_{54}N_7O_{10}$	756.392669	756.39208	-0.78
y7(CB)	1+	$C_{46}H_{63}N_8O_{11}$	903.461083	903.46177	0.76
y8(CB)	1+	$C_{48}H_{66}N_9O_{12}$	960.482547	960.48339	0.88
M-H_2O-NH_3	5+	$C_{257}H_{383}N_{64}O_{76}S_6$	1154.927958	1154.92738	-0.50
M-NH3	5+	$C_{257}H_{386}N_{65}O_{76}S_6$	1158.934486	1158.93514	0.56
				Average	0.11 ppm
				Absolute Average	0.43 ppm
				Standard Deviation	0.52 ppm

Table S3.7 Assignments for Figure 3.26: ECD MS/MS of the C_B produced from radical-induced reduction.

Assignment	Charge State	Elemental Composition	Theoretical Mass (m/z)	Observed Mass (m/z)	Mass Error (ppm)
c3	1+	$C_{18}H_{28}N_5O_4$	378.213581	378.21359	0.02
z4	1+	$C_{19}H_{34}N_4O_7$	430.242202	430.24208	-0.28
a4+H	1+	$C_{22}H_{34}N_6O_5$	462.258520	462.25848	-0.09
c4	1+	$C_{23}H_{36}N_7O_6$	506.272159	506.27205	-0.22
z5	1+	$C_{28}H_{43}N_5O_9$	593.305531	593.30555	0.03
c5	1+	$C_{29}H_{43}N_{10}O_7$	643.331071	643.33110	0.05
a6+H	1+	$C_{34}H_{52}N_{10}O_7$	712.401496	712.40200	0.71
z6	1+	$C_{37}H_{52}N_6O_{10}$	740.373945	740.37495	1.36
c6	1+	$C_{35}H_{54}N_{11}O_8$	756.415135	756.41429	-1.12
z14+H	2+	$C_{75}H_{113}N_{17}O_{20}S_1$	801.902878	801.90199	-1.11
a8+H	1+	$C_{39}H_{60}N_{12}O_9S_1$	872.432145	872.43264	0.57
z15-H	2+	$C_{84}H_{120}N_{18}O_{22}S_1$	882.426717	882.42742	0.80
z7	1+	$C_{46}H_{61}N_7O_{11}$	887.442359	887.44275	0.44
c8	1+	$C_{40}H_{62}N_{13}O_{10}S_1$	916.445784	916.44630	0.56
z16	2+	$C_{90}H_{132}N_{19}O_{23}S_1$	939.472662	939.47290	0.25
z8	1+	$C_{48}H_{64}N_8O_{12}$	944.463823	944.46525	1.51
c17-H	2+	$C_{89}H_{137}N_{24}O_{23}S_1$	970.999909	971.00070	0.81
z17	2+	$C_{93}H_{137}N_{20}O_{24}S_1$	974.991219	974.99128	0.06
a18+H	2+	$C_{93}H_{145}N_{24}O_{23}S_1$	999.031209	999.03129	0.08
z29-2H	3+	$C_{149}H_{224}N_{38}O_{41}S_2$	1088.534541	1088.53320	-1.23
a29	3+	$C_{153}H_{228}N_{39}O_{38}S_2$	1094.551084	1094.54987	-1.11
a10+H	1+	$C_{48}H_{72}N_{16}O_{12}S_1$	1096.523086	1096.52253	-0.51
z9+H	1+	$C_{54}H_{77}N_{12}O_{13}$	1101.572759	1101.57219	-0.52
c21-2H	2+	$C_{104}H_{160}N_{28}O_{29}S_2$	1164.566825	1164.56630	-0.45
z21	2+	$C_{115}H_{171}N_{26}O_{30}S_1$	1214.118211	1214.11788	-0.27
a22+H	2+	$C_{109}H_{172}N_{31}O_{29}S_2$	1221.618386	1221.61763	-0.62
c22-2H	2+	$C_{110}H_{172}N_{32}O_{30}S_2$	1242.617381	1242.61600	-1.11
c24-H	2+	$C_{121}H_{185}N_{34}O_{32}S_2$	1345.166232	1345.16583	-0.30
a25+H	2+	$C_{129}H_{193}N_{34}O_{32}S_2$	1397.197532	1397.19594	-1.14
c25-2H	2+	$C_{130}H_{193}N_{35}O_{33}S_2$	1418.196527	1418.19583	-0.49
z26-H	2+	$C_{135}H_{201}N_{33}O_{36}S_2$	1462.217125	1462.21591	-0.83
c26-H	2+	$C_{139}H_{203}N_{36}O_{35}S_2$	1500.232104	1500.23240	0.20
z27-H	2+	$C_{140}H_{209}N_{35}O_{38}S_2$	1526.246414	1526.24592	-0.32
z28	2+	$C_{144}H_{216}N_{37}O_{40}S_2$	1583.771791	1583.77034	-0.92
				Average	-0.15 ppm
				Absolute Average	0.59 ppm
				Standard Deviation	0.72 ppm

Table S3.8 Assignments for Figure 3.27: ECD MS/MS of the C_B produced from chemical reduction.

Assignment	Charge State	Elemental Composition	Theoretical Mass (m/z)	Observed Mass (m/z)	Mass Error (ppm)
c3	1+	$C_{18}H_{28}N_5O_4$	378.213581	378.21356	-0.06
z.4	1+	$C_{19}H_{35}N_4O_7$	431.250027	431.25003	0.01
c4	1+	$C_{23}H_{36}N_7O_6$	506.272159	506.27170	-0.91
c.20	4+	$C_{99}H_{156}N_{27}O_{26}S_2$	550.778358	550.77851	0.28
z5	1+	$C_{28}H_{43}N_5O_9$	593.305531	593.30497	-0.95
c5	1+	$C_{29}H_{43}N_{10}O_7$	643.331071	643.33113	0.09
z6	1+	$C_{37}H_{52}N_6O_{10}$	740.373945	740.37425	0.41
c6	1+	$C_{35}H_{54}N_{11}O_8$	756.415135	756.41510	-0.05
z7	1+	$C_{46}H_{61}N_7O_{11}$	887.442359	887.44251	0.17
z8	1+	$C_{48}H_{64}N_8O_{12}$	944.463823	944.46325	-0.61
z.28	3+	$C_{144}H_{218}N_{37}O_{40}S_2$	1056.519561	1056.51965	0.08
z.9	1+	$C_{54}H_{77}N_{12}O_{13}$	1101.572759	1101.57297	0.19
c29	3+	$C_{154}H_{231}N_{40}O_{39}S_2$	1109.558238	1109.55869	0.41
c.21	2+	$C_{104}H_{161}N_{28}O_{29}S_2$	1165.070738	1165.07054	-0.17
z.10	1+	$C_{59}H_{84}N_{13}O_{16}$	1230.615353	1230.61542	0.05
c22	2+	$C_{110}H_{174}N_{32}O_{30}S_2$	1243.625206	1243.62487	-0.27
c23	2+	$C_{112}H_{177}N_{33}O_{31}S_2$	1272.135938	1272.13533	-0.48
c.24	2+	$C_{121}H_{185}N_{34}O_{32}S_2$	1345.166232	1345.16550	-0.54
c.25	2+	$C_{130}H_{194}N_{35}O_{33}S_2$	1418.700439	1418.70004	-0.28
c.26	2+	$C_{139}H_{203}N_{36}O_{35}S_2$	1500.232104	1500.23269	0.39
				Average	-0.11 ppm
				Absolute Average	0.32 ppm
				Standard Deviation	0.41 ppm

Table S3.9 Assignments for Figure 3.28: CAD MS/MS of the C_B produced from chemical reduction.

Assignment	Charge State	Elemental Composition	Theoretical Mass (m/z)	Observed Mass (m/z)	Mass Error (ppm)
b2-CO	1+	$C_{13}H_{19}N_2O_1$	219.149189	219.14921	0.10
b2	1+	$C_{14}H_{19}N_2O_2$	247.144104	247.14408	-0.10
y2	1+	$C_{10}H_{22}N_3O_4$	248.160483	248.16046	-0.09
y3-H_2O	1+	$C_{15}H_{27}N_4O_4$	327.202682	327.20270	0.06
y3-NH_3	1+	$C_{15}H_{26}N_3O_5$	328.186698	328.18659	-0.33
y3	1+	$C_{15}H_{29}N_4O_5$	345.213247	345.21328	0.10
b3	1+	$C_{18}H_{25}N_4O_4$	361.187032	361.18701	-0.06
y4-H_2O	1+	$C_{19}H_{34}N_5O_6$	428.250361	428.25040	0.09
y4	1+	$C_{19}H_{36}N_5O_7$	446.260926	446.26086	-0.15
b4-NH_3	1+	$C_{23}H_{30}N_5O_6$	472.219061	472.21881	-0.53
b4	1+	$C_{23}H_{33}N_6O_6$	489.245610	489.24546	-0.31
b10-CO	2+	$C_{48}H_{72}N_{16}O_{12}S_1$	548.261268	548.26120	-0.12
y9-H_2O	2+	$C_{54}H_{77}N_{13}O_{12}$	549.790185	549.79046	0.50
b10-H_2O	2+	$C_{49}H_{70}N_{16}O_{12}S_1$	553.253444	553.25346	0.03
y9	2+	$C_{54}H_{79}N_{13}O_{13}$	558.795467	558.79533	-0.25
b10	2+	$C_{49}H_{72}N_{16}O_{13}S_1$	562.258726	562.25872	-0.01
b5-CO_2	1+	$C_{28}H_{40}N_9O_5$	582.314692	582.31443	-0.45
y5-H_2O	1+	$C_{28}H_{43}N_6O_8$	591.313690	591.31348	-0.36
b5-CO	1+	$C_{28}H_{40}N_9O_6$	598.309607	598.30958	-0.05
b11-CO	2+	$C_{54}H_{83}N_{17}O_{13}S_1$	604.803300	604.80334	0.07
b5-H_2O	1+	$C_{29}H_{38}N_9O_6$	608.293957	608.29412	0.27
y5	1+	$C_{28}H_{45}N_6O_9$	609.324255	609.32421	-0.07
b11-NH_3	2+	$C_{55}H_{80}N_{16}O_{14}S_1$	610.287484	610.28759	0.17
y10-H_2O	2+	$C_{59}H_{84}N_{14}O_{15}$	614.311482	614.31126	-0.36
b11	2+	$C_{55}H_{83}N_{17}O_{14}S_1$	618.800758	618.80073	-0.05
y10	2+	$C_{59}H_{86}N_{14}O_{16}$	623.316764	623.31630	-0.74
b5	1+	$C_{29}H_{40}N_9O_7$	626.304522	626.30455	0.04
y11-H_2O	2+	$C_{61}H_{87}N_{15}O_{16}$	642.822214	642.82189	-0.50
y11	2+	$C_{61}H_{89}N_{15}O_{17}$	651.827496	651.82757	0.11
b12-CO	2+	$C_{59}H_{92}N_{18}O_{14}S_1$	654.337507	654.33763	0.19
b12-H_2O	2+	$C_{60}H_{90}N_{18}O_{14}S_1$	659.329682	659.33029	0.92
b12-NH_3	2+	$C_{60}H_{89}N_{17}O_{15}S_1$	659.821691	659.82183	0.21
b12	2+	$C_{60}H_{92}N_{18}O_{15}S_1$	668.334965	668.33527	0.46
y12	2+	$C_{64}H_{94}N_{16}O_{18}S_1$	703.332089	703.33223	0.20
b6-CO	1+	$C_{34}H_{51}N_{10}O_7$	711.393671	711.39321	-0.65
b13-CO	2+	$C_{64}H_{99}N_{19}O_{17}S_1$	718.858805	718.85833	-0.66
b13-NH_3	2+	$C_{65}H_{96}N_{18}O_{18}S_1$	724.342988	724.34243	-0.77

b13	2+	$C_{65}H_{99}N_{19}O_{18}S_1$	732.856262	732.85655	0.39
y26	4+	$C_{135}H_{206}N_{34}O_{36}S_2$	735.868838	735.86871	-0.17
b6	1+	$C_{35}H_{51}N_{10}O_8$	739.388586	739.38877	0.25
y13-H_2O	2+	$C_{69}H_{101}N_{17}O_{18}S_1$	743.861013	743.86166	0.87
y13	2+	$C_{69}H_{103}N_{17}O_{19}S_1$	752.866295	752.86616	-0.18
y6	1+	$C_{37}H_{54}N_7O_{10}$	756.392669	756.39277	0.13
b14-NH_3	2+	$C_{68}H_{101}N_{19}O_{19}S_1$	759.861545	759.86133	-0.28
y27-H_2O	4+	$C_{140}H_{212}N_{36}O_{37}S_2$	763.380841	763.38021	-0.83
b14	2+	$C_{68}H_{104}N_{20}O_{19}S_1$	768.374819	768.37472	-0.13
y28	4+	$C_{144}H_{220}N_{38}O_{40}S_2$	796.394214	796.39409	-0.16
y21	3+	$C_{115}H_{174}N_{27}O_{30}S_1$	815.087474	815.08781	0.41
y29	4+	$C_{149}H_{229}N_{39}O_{41}S_2$	821.161318	821.16124	-0.09
b15	2+	$C_{74}H_{115}N_{21}O_{20}S_1$	824.916851	824.91722	0.45
b23	3+	$C_{112}H_{175}N_{32}O_{31}S_2$	842.750868	842.75092	0.06
y23	3+	$C_{120}H_{182}N_{29}O_{33}S_1$	863.105305	863.10596	0.76
y15-H_2O	2+	$C_{84}H_{121}N_{19}O_{21}S_1$	881.934710	881.93473	0.02
b8	1+	$C_{40}H_{59}N_{12}O_{10}S_1$	899.419235	899.41952	0.32
b16	2+	$C_{83}H_{124}N_{22}O_{22}S_1$	906.448515	906.44796	-0.61
y8	1+	$C_{48}H_{66}N_9O_{12}$	960.482547	960.48315	0.63
b17	2+	$C_{89}H_{135}N_{23}O_{23}S_1$	962.990548	962.99139	0.87
b9-H_2O	1+	$C_{43}H_{62}N_{13}O_{11}S_1$	968.440699	968.44034	-0.37
b28-NH_3	3+	$C_{148}H_{213}N_{36}O_{38}S_2$	1055.508885	1055.50943	0.52
y19-H_2O	2+	$C_{103}H_{153}N_{23}O_{27}S_1$	1088.050802	1088.05099	0.17
y19	2+	$C_{103}H_{155}N_{23}O_{28}S_1$	1097.056085	1097.05628	0.18
y20	2+	$C_{109}H_{166}N_{24}O_{29}S_1$	1153.598117	1153.59720	-0.79
b21	2+	$C_{104}H_{159}N_{27}O_{29}S_2$	1157.061376	1157.06055	-0.71
b22	2+	$C_{110}H_{171}N_{31}O_{30}S_2$	1235.111931	1235.11234	0.33
b23	2+	$C_{112}H_{174}N_{32}O_{31}S_2$	1263.622663	1263.62164	-0.81
y24-NH_3	2+	$C_{123}H_{183}N_{29}O_{34}S_2$	1337.145637	1337.14504	-0.45
y24	2+	$C_{123}H_{186}N_{30}O_{34}S_2$	1345.658912	1345.65975	0.62
b25	2+	$C_{130}H_{192}N_{34}O_{33}S_2$	1410.691077	1410.69205	0.69
				Average	-0.01 ppm
				Absolute Average	0.34 ppm
				Standard Deviation	0.44 ppm

Chapter 4

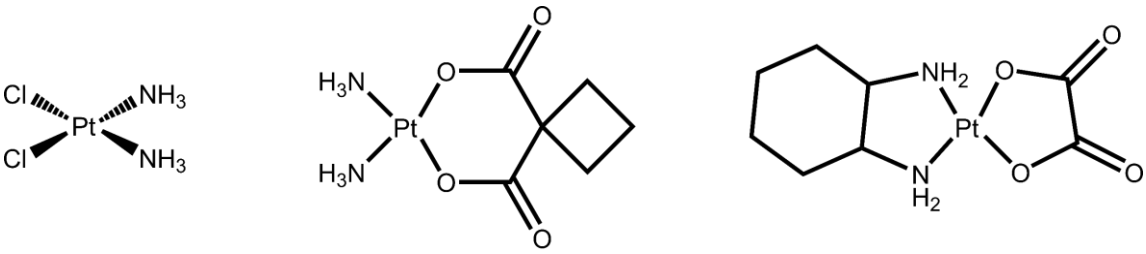
Organoiridium Photosensitisers Induce Specific Oxidative Attack on Proteins within Cancer Cells

The photocatalytic octahedral iridium-(III)-based anticancer complexes, $[Ir(C,N)_2(S,S)]$ (Series 1) and $[Ir(C,N)_2(O,O)]^+$ (Series 2), where C,N = phenylquinoline, and O,O = a diketonate, S,S = a dithione chelating ligands, show the ability of generating singlet oxygen ($^1O_2$) upon irradiation with visible blue light (463 nm), and oxidise the model peptide, [Lys^3]-Bombesin (L3BBS). Proteomics studies of human lung cancer cells (A549) revealed that photoactivation of the iridium complex caused selective oxidative damage to specific histidine residues in the key cellular proteins heat shock protein 70 kDa, which is involved in protein folding and protection of cells from stress, and aldose reductase, which plays a key role in hyperglycemia-induced oxidative stress. The oxidative stress induced by iridium photosensitisers during photoactivation increased the levels of enzymes involved in the glycolytic pathway.

Iridium-(III)-based photosensitisers were synthesised and their IC_{50} values were studied by Dr. Pingyu Zhang. The drug treated A549 lung cancer cell cultures were prepared by Dr. Pingyu Zhang and Dr. Huaiyi Huang. Cell lysis and nLC studies were carried out with the assistance of Dr. Yuko P. Y. Lam. Sample preparation and all data were produced, analysed and written by the thesis Author. A significant part of this chapter was published in Angewandte Chemie International Edition, as a paper entitled "Organoiridium Photosensitizers Induce Specific Oxidative Attack on Proteins within Cancer Cells" https://onlinelibrary.wiley.com/doi/full/10.1002/anie.201709082

Introduction

Metal-based chemotherapies have been widely used for many years around the world. Platinum-based compounds, such as cisplatin, oxaliplatin, and carboplatin (Scheme 1) are the most commonly used metal-based anticancer agents to treat different types of cancers. However platinum resistance is now a common problem.[1] Every metallodrug has its own mechanism of action (MoA). Cisplatin's mechanism is centred around reacting with the DNA of the cancerous cells, disrupting its structure, resulting in failure of transcription, i.e. anti-proliferation. However, studies have shown 90-98% of cisplatin becomes protein-bound after just one day of injection, meaning most of the drug cannot reach the desired target (DNA of the cancerous cells),[2,3] vastly influencing the effectiveness of the treatment, and inducing undesired side effects to the patients. Therefore, there is an essential need to develop different metal-based anticancer drugs, with different mechanisms of action to attack cancer cells. Different metallodrugs, such as iridium,[4,5] rhodium,[6] platinum,[7] osmium,[8,9] gold,[10] ruthenium[11], have been designed to possess different properties. Interest in the use of luminescent transition metal complexes for various application has increased enormously over the past several decades, particularly as photosensitisers for photodynamic therapy (PDT).[12–14] PDT has been widely used in the last 40 years, PDT involves absorption of light by a photosensitiser promoting formation of a triple state that can polarise the spins of the ground-state $^3O_2$ to generate highly toxic singlet oxygen ($^1O_2$). However, selectivity is also an important criterion for being an effective anticancer drug, as these drugs are often cytotoxic, if selectivity is low, healthy cells could be affected or even killed too. To enhance selectivity of a metallodrug, pro-drugs (inactive drugs) are often designed to possess low cytotoxicity and only become highly cytotoxic when activated, usually by the change in chemical environment, such as pH or salt concentration, or even activated with light, also known as photoactivatable chemotherapy (PACT).[15]



cisplatin carboplatin oxaliplatin

Scheme 4.1 Platinum-based anticancer drugs.

Different analytical techniques have been utilised to study the interaction between metallodrugs and biomolecules, such as ^{1}H-NMR or x-ray crystallography.[16,17] These techniques are often used to provide structural information and sometimes able to identify and locate the modifications on biomolecules. However, due to the inherent lack of separation ability, the capabilities of these analytical techniques are often limited. Other techniques which possess the ability of handling mixtures of products/adducts are required to study these metallodrug-biomolecules systems more comprehensively. Mass spectrometry (MS) is increasing being used to study biological systems, due to its high sensitivity and ability to cope with complex samples. Electrospray ionisation (ESI),[18] and nano electrospray ionisation (nESI),[19] are the most commonly used soft ionisation techniques to generate multiply-charged species, which allow to study large biomolecules in the lower mass-to-charge ratio (*m/z*) region. MS can also be coupled to hyphenated techniques, such as liquid chromatography (LC) or gas chromatography (GC), which allows separation prior to MS analysis and simplify the complexity of the sample. MS also has the ability of separating species of interest, often using a quadrupole, tandem mass spectrometry (MS/MS) experiments which can be performed to fragment the separated ions, obtaining structural/sequence informative fragments. An extensive array of fragmentation techniques has been developed, including collisionally activated dissociation (CAD),[20] electron capture dissociation (ECD),[21] infrared multi-photon dissociation (IRMPD),[22] ultraviolet photodissociation (UVPD)[23] etc.

Fourier transform-ion cyclotron resonance mass spectrometry (FT-ICR MS) uses both electric and magnetic fields to trap species of interest for detection. FT-ICR MS is

well-known for its ultra-high resolving power and mass accuracy, which is not comparable by other mass spectrometers. With the capability of isolating ions of interest using a quadrupole, applying collisional activation, and electron/photon-based dissociation techniques upon ions of interests, FT-ICR MS provides the largest array of fragmentation techniques and highest mass accuracy of any mass spectrometer, with mass errors into sub parts-per-million (ppm) and even parts-per-billion (ppb) ranges.

Herein, the interaction and MoA of [Lys^3]-Bombesin (L3BBS) and iridium-(III)-based complexes (PZ1-PZ6) was studied using ultra-high resolution FT-ICR MS together with various MS/MS techniques to characterise the oxidation sites on the peptide. Reversed phase (RP) nano-LC (nLC) was coupled to the FT-ICR MS to study the effect of the most potent iridium complex (PZ6) on lung cancer cell line A549 and perform quantitative MS using label-free quantification.

Experimental

Materials

[Lys^3]-Bombesin (L3BBS), chicken lysozyme, formic acid, ammonium bicarbonate (AmmBic), dithiothreitol (DTT), iodoacetamide (IAA), and trypsin were all purchased from Sigma Aldrich (Dorset, England). Human lung carcinoma (A549) cell line was purchased from the European Collection of Authenticated Cell Cultures (ECACC) (Salisbury, UK). LCMS grade acetonitrile (MeCN) was purchased from VWR (Lutterworth, UK). LCMS grade methanol (CH_3OH), hypersep C18 RP cartridges were purchased from Thermo Fisher Scientific (Waltham, UK). Ultra-pure water (18.2 $M\Omega cm^{-1}$) was obtained from a Milli-Q UV III water purification system (Milli-Q, Hertfordshire, UK). The dinuclear iridium complex [$(C,N)_4Ir_2Cl_2$] was prepared by Dr. Pingyu Zhang according to literature methods reported below.[24]

Synthesis of the iridium-(III)-based complexes (PZ1-PZ3)

A mixture containing the dithione (0.015 mmol), Ag_2O (0.15 mmol) and the corresponding iridium dimer (0.07 mmol) in CH_2Cl_2 (50 mL) was heated in a sealed tube overnight at 348 K. After cooling to ambient temperature, the mixture was filtered through celite and washed with CH_2Cl_2. The filtrate was concentrated under reduced

pressure and purified on a silica column by Biotage flash chromatography with CH_2Cl_2:CH_3OH (v/v 80:20) as eluant.

Synthesis of the iridium-(III)-based complexes (PZ4-PZ6)

A mixture containing the diketonate (0.015 mmol) and the corresponding iridium dimer (0.07 mmol) in CH_2Cl_2 (50 mL) was heated in a sealed tube for 2 hours at 348 K. After cooling to ambient temperature, the mixture was filtered through celite and washed with CH_2Cl_2. The filtrate was concentrated under reduced pressure and purified on a silica column by Biotage flash chromatography with CH_2Cl_2:CH_3OH (v/v 95:5) as eluant.

FT-ICR MS study of PZx-L3BBS reactions

Aqueous solutions of L3BBS (1 mM) and iridium-(III)-based complexes (PZ1-PZ6, 250 µM, in 50:50 H_2O:MeCN) were mixed in 1:2 and 1:5 (peptide:drug) molar ratios. The sample incubation and irradiation conditions mirrored the IC_{50} testing conditions for each compound. Briefly samples were incubated for 1 h or 2 h (for PZ1-PZ3, and PZ4-PZ6 respectively) at 348 K. Followed by 5 min of irradiation with 463 nm visible blue light, and further incubation of 47 h or 46 h for PZ1-PZ3, and PZ4-PZ6, respectively. All sample preparation, incubations, and mass spectrometry analysis were conducted in darkness or with very low levels of red light to avoid activation of the complexes. Samples were then diluted with ultra-pure water to 5 µM total concentration for FT-ICR MS analysis (Figure 4.1). All samples were analysed using nESI for increased sensitivity and lower the sample consumption compared to conventional ESI. All MS experiments were carried out using a SolariX FT-ICR mass spectrometer, fitted with a 12 tesla actively shielded magnet (Bruker Daltonik GmbH, Bremen, Germany). All samples were sprayed in positive ion mode. For MS experiments, ions were accumulated for 0.5 s in the hexapole-based collision cell before transferred to the infinity cell for excitation and detection.

For ECD MS/MS experiments, ions of interest were isolated using the front-end quadrupole using a wide isolation window (30-50 *m/z*), then accumulated in the hexapole for 4 s. The isolated ion packets of interest were then transferred into the

infinity cell. In cell isolation, Multi-Correlated Harmonic Excitation Fields (Multi-CHEF),[25] was also used to further isolate the ions of interest, with notch width of 0.5 *m/z*, and isolation power of 50%, to achieve a cleaner isolation. The isolated ions were then irradiated with 1.2 eV electrons, with pulse length of 0.15 s, generated from a 1.5 A hollow cathode dispenser *via* an extraction lens held at 18 V.

All spectra were internally calibrated using a quadratic calibration function, manually interpreted, and assigned *via* DataAnalysis v4.2 (Bruker Daltonik GmbH, Bremen, Germany).

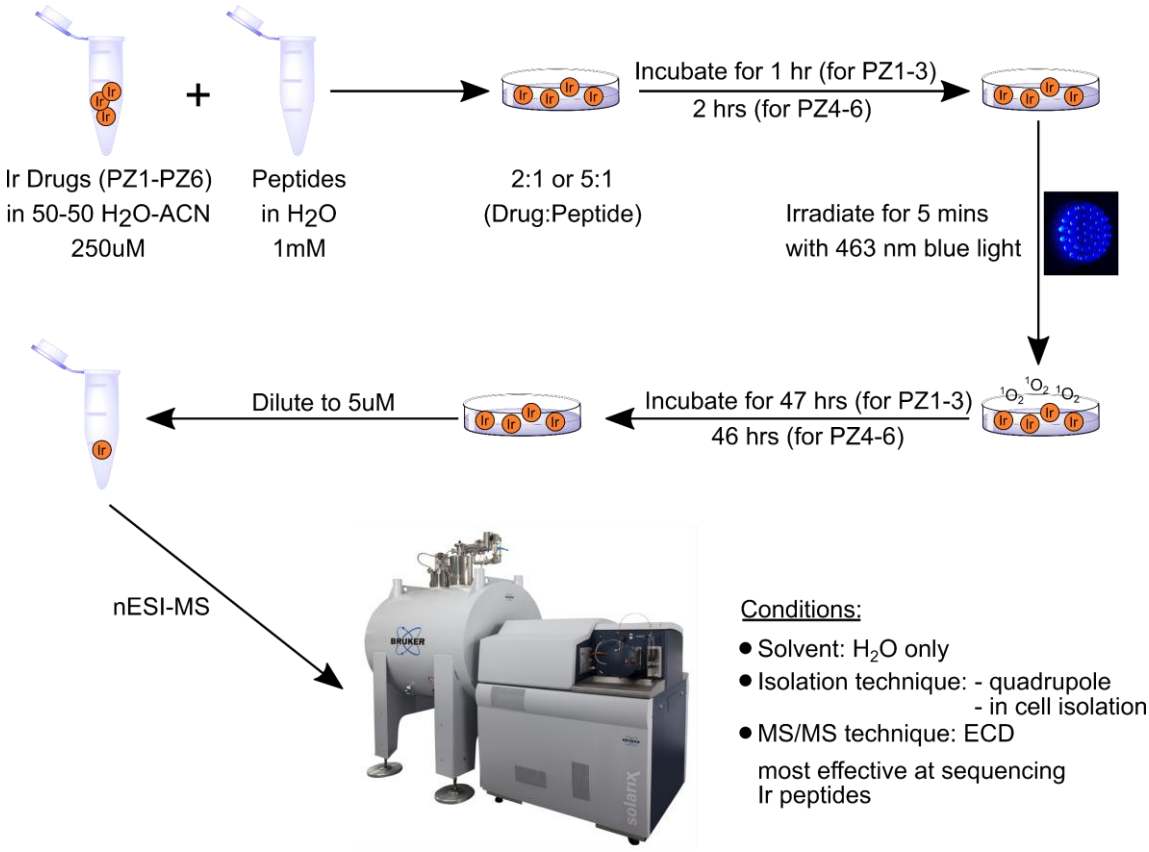


Figure 4.1 FT-ICR MS study of PZx-L3BBS reaction.

nLC-FT-ICR MS/MS of the PZ6-A549 lung cancer cell line

Proteomics for Target Site Identification

Approximately $1x10^9$ A549 lung cancer cells were treated with PZ6 (10 μM) in the dark or with 10 min of blue LED irradiation (463 nm). All cells were harvested and lysed by sonicating with glass beads at 277 K in a Tris-buffer solution (50 mM Tris-HCl with 1% glycerol, pH 7.5). The samples were sonicated for 30 s and cooled for 5 min, the sonication process being repeated 10 times. All samples were centrifuged for 30 min at 5,000 g at 277 K to remove the pellets. The supernatant from each sample was then quantified against a standard calibration curve of bovine serum albumin (BSA) using Bradford Reagent (Sigma Aldrich Company Ltd; Dorset, England). The concentrations of proteins in these samples were then calculated.

The above samples were diluted using 100 mM AmmBic solution (Sigma Aldrich Company Ltd; Dorset, England); the disulfide bonds were reduced with 50 mM DTT (Sigma Aldrich Company Ltd; Dorset, England) and the free cysteine thiols were alkylated with 100 mM IAA (Sigma Aldrich Company Ltd; Dorset, England). The alkylated lysate was then digested with trypsin (Sigma Aldrich Company Ltd; Dorset, England) by incubating the mixture for 16 h at 310 K. The digested peptides were desalted using C18 RP cartridges (Thermo Fisher Scientific, Waltham, UK), washed twice with Milli-Q water and then eluted with 80% MeCN/aqueous solution. Desalted samples were dried down with Speed Vac concentrator (Thermo Fisher Scientific, Waltham, UK) and re-dissolved in Milli-Q water into concentration of 1 mg/mL.

nLC separations were achieved using an EASY nano-LC II system (Proxeon, Hemel Hempstead, UK) with a in-home-made 18 cm, C18 RP nano capillary column (75 μm, 5 μm particle size, pore size of 300 Å) and a 3 cm C18 RP pre-column (150 μm, 5 μm particle size, pore size of 300 Å). Separation of tryptic digested cell was achieved using an acidified H_2O/MeCN gradient from 5% MeCN to 30% MeCN over 120 min, acidified with formic acid, followed by a second gradient of 15 min from 30% MeCN to 80% MeCN. Finally, with a 35-min wash of 80% MeCN at a constant 600 nL/min flow rate.

nLC-MS/MS experiments were carried out by coupling EASY-nLC to SolariX FT-ICR MS, *via* a custom nano-spray source utilising glass nanospray tips, 360 µm tip O.D., 50 µm tip I.D., and 8 µm orifice I.D. (New Objective, MA, USA). Singly-charged or unknown charged species were excluded automatically from the MS/MS list, only the most intense multiply-charged ions from each scan were isolated in the quadrupole and accumulated for 1.2 s in the hexapole collision cell for CAD MS/MS experiments. 1.6 µg of each sample was injected for each nLC-MS/MS run. All nLC-MS/MS experiments were performed with 5 technical replicates for consistency and reproducibility purposes, and more importantly, to increase the reliability of the data sets.

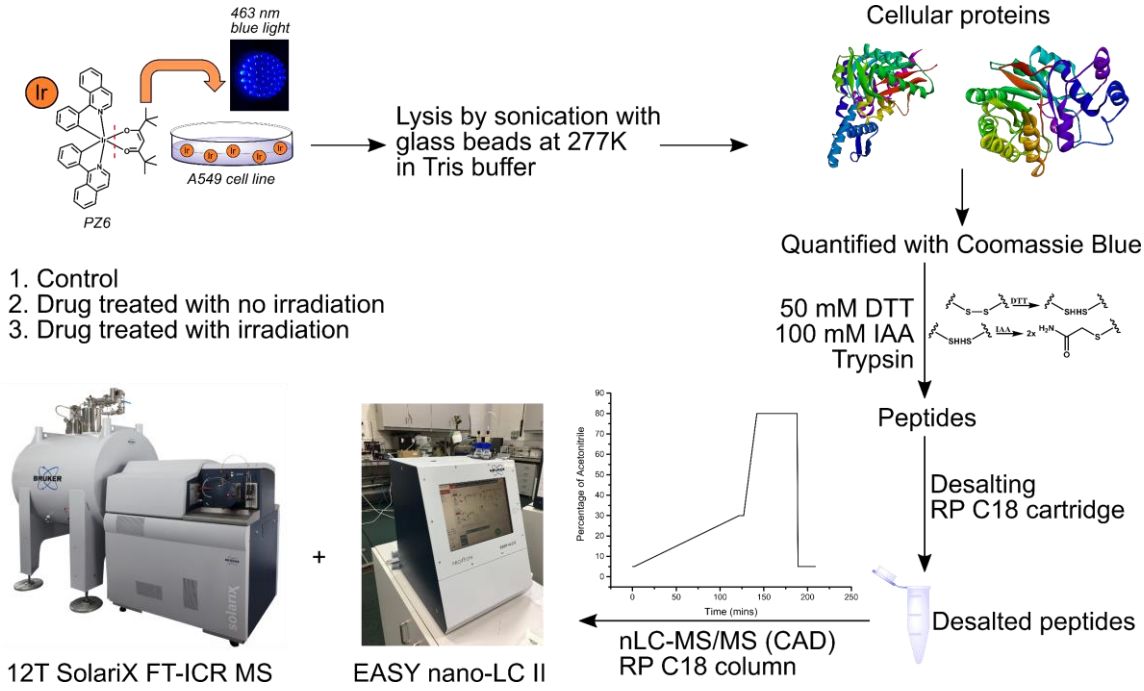


Figure 4.2 nLC-FT-ICR MS/MS of the PZ6-A549 lung cancer cell line.

Mass lists were generated using the FTMS peak picking algorithm in DataAnalysis v4.2 (Bruker Daltonik GmbH, Bremen, Germany) from nLC-MS/MS data sets. Then the mass lists were searched against the Mascot database search engine and all identified proteins were filtered with 1% false discovery rate (FDR). Each oxidised peptide picked out by the search engine was manually analysed from the raw nLC-MS/MS data, and any false positive results were critically eliminated. The reduced

datasets (after elimination), were then grouped together, and oxidised peptides that were detected three times or more out of the five replicates of data were further selected for quantification.

Singlet oxygen-induced oxidation of proteins is complicated as discussed in several reviews being the result of direct attack on amino acids and secondary attack by initially formed peroxides.[26,27] On the basis of previous reports, the following modifications were included in the search:

Variable modifications: methionine to sulfoxide (+O), methionine to sulfone (+2O), histidine to 2-oxo-histidine (+O), tryptophan to kynurenine (+O_2-CO), tryptophan to 3-hydroxykynurenine (+O_2-CO+O), tryptophan to N-formylkynurenine (+2O), cysteine to sulfenic acid (+O), cysteine to sulfinic acid (+2O), cysteine to sulfonic acid (+3O), phenylalanine to dihydroxyphenylalanine (+2O), tyrosine to dihydroxyphenylalanine (+O).

In addition, a fixed carbamidomethyl modification of cysteine (from alkylation) and variable deamidation modifications of asparagine, glutamine, and phosphorylation of serine, threonine, and tyrosine were included.

Quantification of peptides was achieved by spiking with a lysozyme tryptic digest in 1:6 lysozyme peptide:cell digest volume ratio (lysozyme peptide 0.01 µg/µL, cell digest 0.2 µg/µL) into each sample during the nLC-MS/MS runs and by determining the ratio between the cell peptide of interest and the specific lysozyme peptide (FESNFNTQATNR, 714.8365 *m/z*). The differences between the ratios from different reaction conditions (control, drug treated with no irradiation, or drug treated with irradiation) were compared. T-tests and p values were determined as measures of the level of significance between different samples.

Pathway Analysis

Peptides that were detected 3 times or more out of the 5 replicates of data sets were selected for quantification of pathway analysis. The protein entry lists were imported and cross-matched with the Database for Annotation, Visualisation and Integrated Discovery (DAVID) Bioinformation Database,[28] then a list of related

pathways was generated and exported by the Kyoto Encyclopedia of Genes and Genomes (KEGG) database.[29–31] Glycolysis, which was found to have the most number of protein counts and lowest Fisher Exact p-values, were chosen for quantification using the methodology described above. The weighted area ratios were calculated using the equation shown below for proteins with multiple identified peptides. This provided a more comprehensive measure of the change in abundance under different conditions.

$$Average\ Fold\ of\ Change = \sum(F \times \frac{A}{\sum A}) \qquad \text{[Eq. 1]}$$

where F is the fold of change of a peptide, A is the average area ratio among the 5 replicates of data sets.

Results and Discussion

Two series of iridium-(III)-based photosensitisers were designed to generate highly toxic singlet oxygen ($^1O_2$) upon irradiation with visible blue light (463 nm). Series one (PZ1-PZ3) ($[Ir(C,N)_2(S,S)]^+$) was designed to retain the dithione ligands (shown in blue in Figure 3) upon irradiation, while series two (PZ4-PZ6) ($[Ir(C,N)_2(O,O)]$) was designed to dissociate the diketonate ligands (shown in green in Figure 3) upon irradiation.

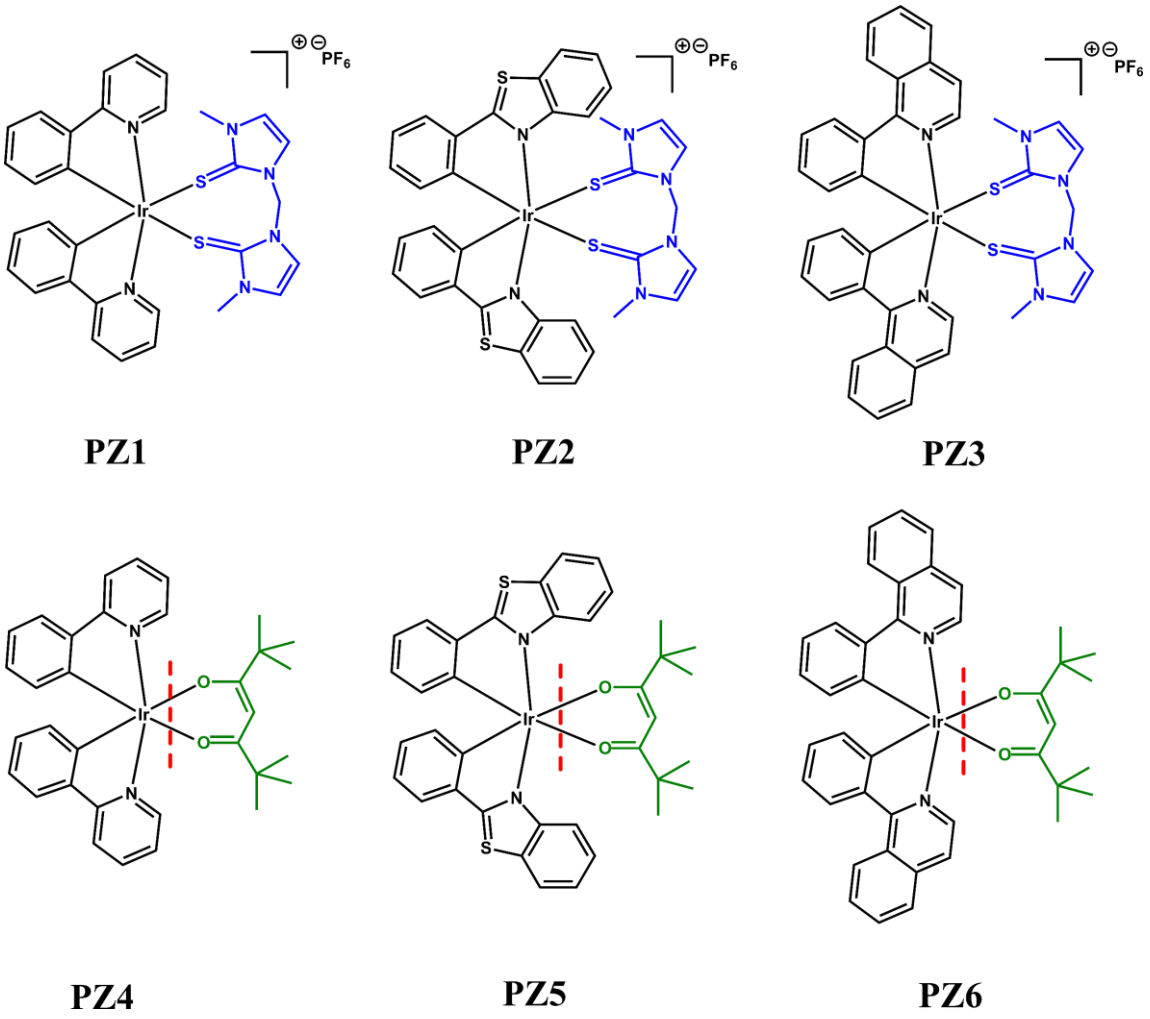


Figure 4.3 Structure of Ir(III)-based photosensitsers, PZ1-PZ6. Dithione ligands (blue) on PZ1-PZ3 are retained upon irradiation, diketonate ligands (green) on PZ4-PZ6 are dissociated upon irradiation (dissociation line in red).

To study the effectiveness of a drug, it is important to investigate a drug's MoA. These iridium-based photosensitisers were designed to release $^1O_2$ upon irradiation, which can oxidise the species in the surrounding environment. Proteins are a major target for oxidants as a result of their high abundance in biological systems. Oxidation can occur at both the protein backbone and on the amino acid side chains. Some side chain oxidation products, including methionine sulfoxide, can be employed as sensitive, specific markers of oxidative damage. Most protein damage is non-repairable, and leads to a loss of specificity of protein structure and function.[32] Some oxidation products are diagnostic for investigating whether the oxidation event is induced by one kind of

reactive oxygen species (ROS) or another. For example, different oxidation products of tryptophan can be used to diagnose whether the oxidation event is induced by hydroxyl radicals (•OH) or $^1O_2$ (Figure 4.4). Figure 4.4 below shows the possible oxidation products of amino acids and the corresponding mass shifts (Δm). MS with sufficient resolving power and mass accuracy, such as FT-ICR MS is able to identify the species using these mass shifts and provide a better understanding of the MoA of a drug.

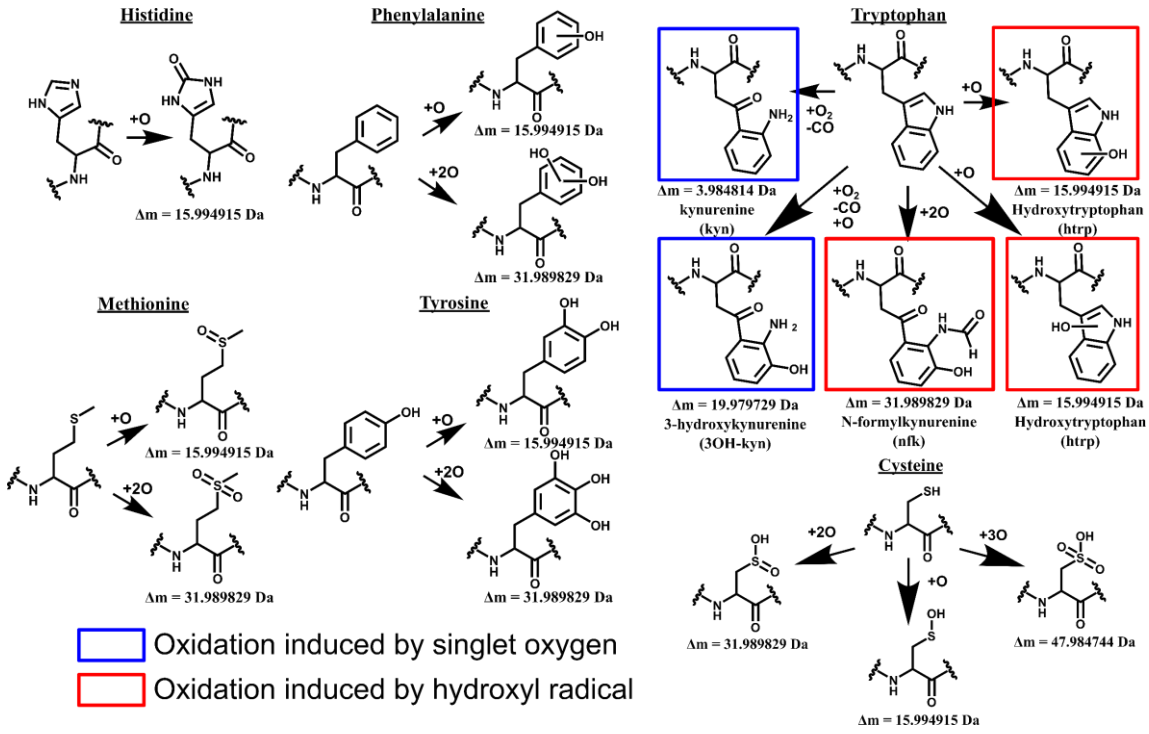


Figure 4.4 Oxidation products of amino acids. $^1O_2$ induced oxidation products of tryptophan are highlighted in blue and hydroxyl radical induced oxidation products of tryptophan are highlighted in red.

Bombesin is a known tumour marker peptide for small cell lung carcinoma, neuroblastoma, and gastric cancer,[33] and therefore was chosen to be a model peptide to study the MoA of the iridium-based photosensitisers. Figures 5 and 6 show the full mass spectrum of 2:1 (drug:peptide) reaction ratio of PZ1:L3BBS and 2:1 reaction ratio of PZ3-L3BBS analysed using nESI-FT-ICR MS. Both spectra showed similar results, a whole array of oxidised L3BBS was observed in the mass spectrum. Apart from the

oxidised products, interestingly, the drugs were found to bind to L3BBS directly, as well as the oxidised forms of L3BBS, $[PZ1\text{-}(S,S)+L3BBS+H]^{2+}$, $[PZ1\text{-}(S,S)+L3BBS+O+H]^{2+}$, and $[PZ1\text{-}(S,S)+L3BBS+2O+H]^{2+}$. However, it was surprising to observe the loss of dithione ligands in these products, as series one (PZ1-PZ3) was designed to retain the dithione ligands upon irradiation. This may indicate the Ir-S bonds are not as unreactive as previously expected. A series of oxidised products were also observed and with the aid of ultra-high resolution mass spectrometry and the ability to obtain high mass accuracies. The oxidised products were assigned according to the mass shifts from the unmodified doubly charged L3BBS species (796.411734 *m/z*). There are a few oxidised products worthy of a more detailed investigation into the loss of ($-CO+O_2$, $\Delta m = 3.984814$ Da). Such modification was seen in $^1O_2$-induced tryptophan oxidation in previous studies (formation of kynurenine, or possibly 3-hydroxykynurenine),[13] which can provide indirect evidence of the generation of $^1O_2$ during the irradiation process.

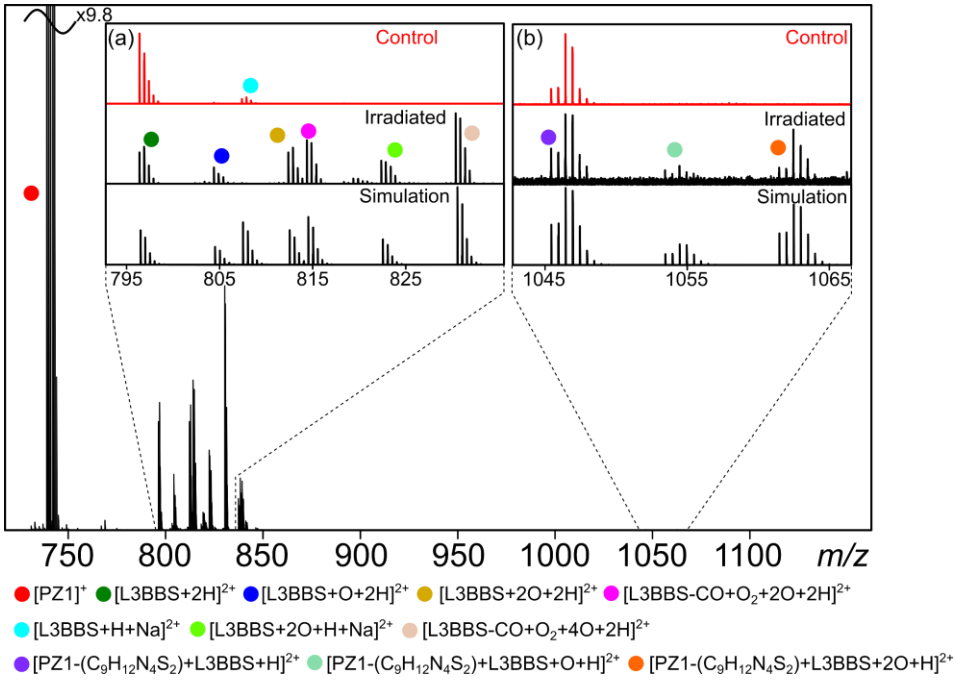


Figure 4.5 Full MS of PZ1-L3BBS (2:1) by nESI. Inset (a) shows the formation of oxidised products. Inset (b) shows the formation of iridium-bound products. Identities of oxidised products and iridium-bound products are listed below the figure.

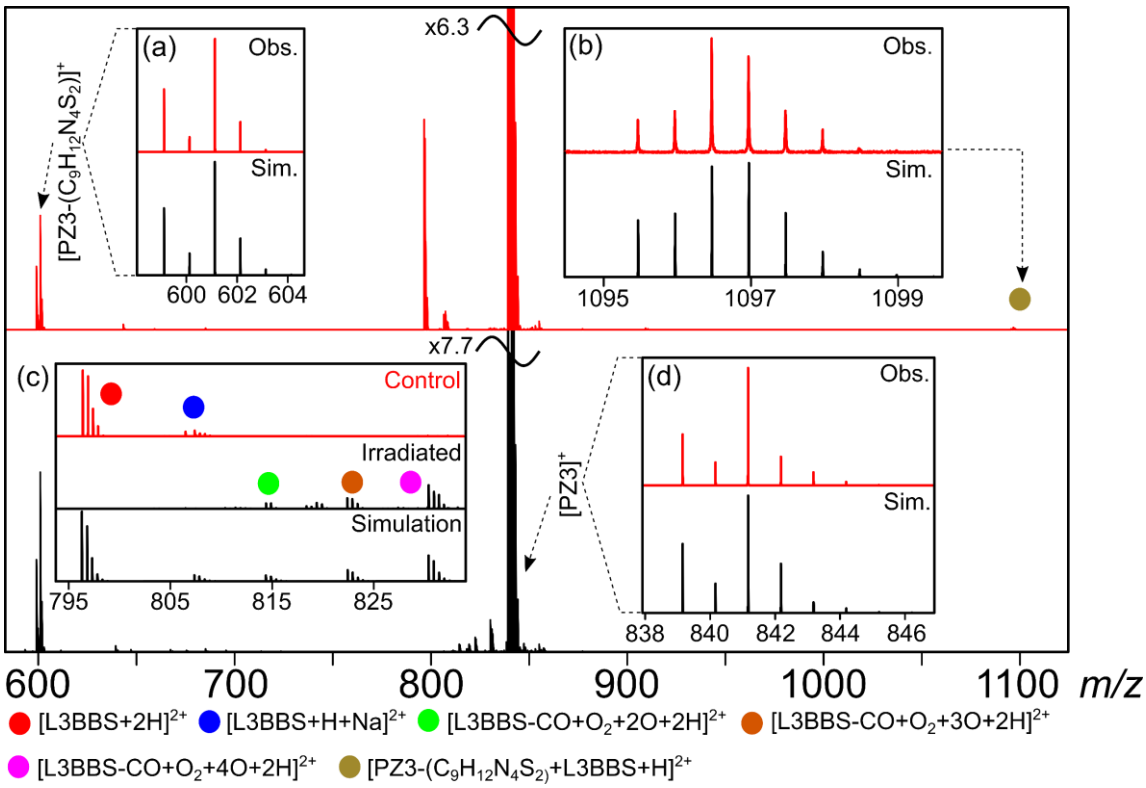


Figure 4.6 Full MS of PZ3-L3BBS (2:1) by nESI. Inset (a) shows the simulation of PZ3 with the loss of dithione ligand. Inset (b) shows the formation of iridium-bound product. Inset (c) shows the formation of oxidised products. Inset (d) shows the intact PZ3 complex. Identities of oxidised products and iridium-bound products are listed below the figure.

Like PZ1 and PZ3, when PZ4 and PZ6 reacted with L3BBS (Figures 4.7 and 4.8), a series of oxidation products were generated as well as the iridium-bound peptides. The diketonate was dissociated as expected with how this series of drug was designed. More importantly, the formation of kynurenine (mass shift of 3.984814 Da) was also observed, indicating the generation of $^1O_2$ in the reaction, which is often observed in other PDT.[14]

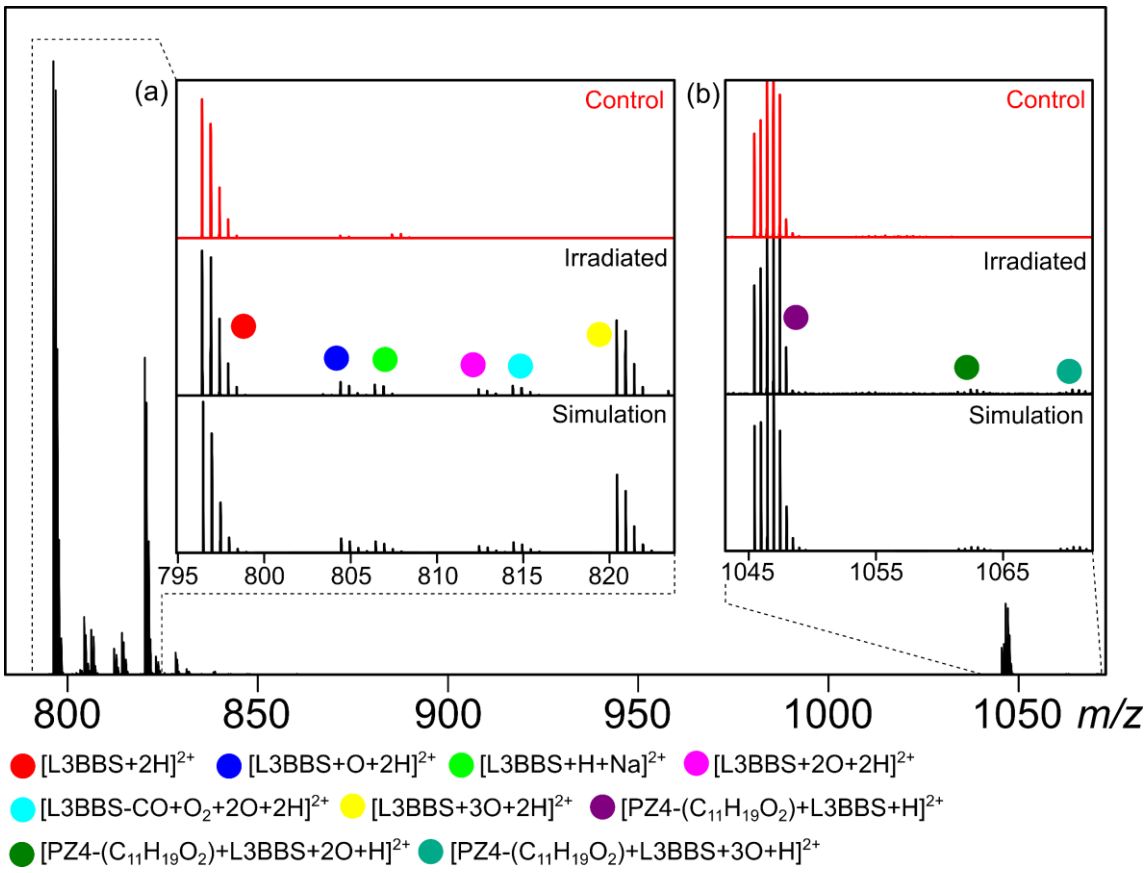


Figure 4.7 Full MS of PZ4-L3BBS (2:1) by nESI. Inset (a) shows the formation of oxidised products. Inset (b) shows the formation of iridium-bound products. Identities of oxidised products and iridium-bound products are listed below the figure.

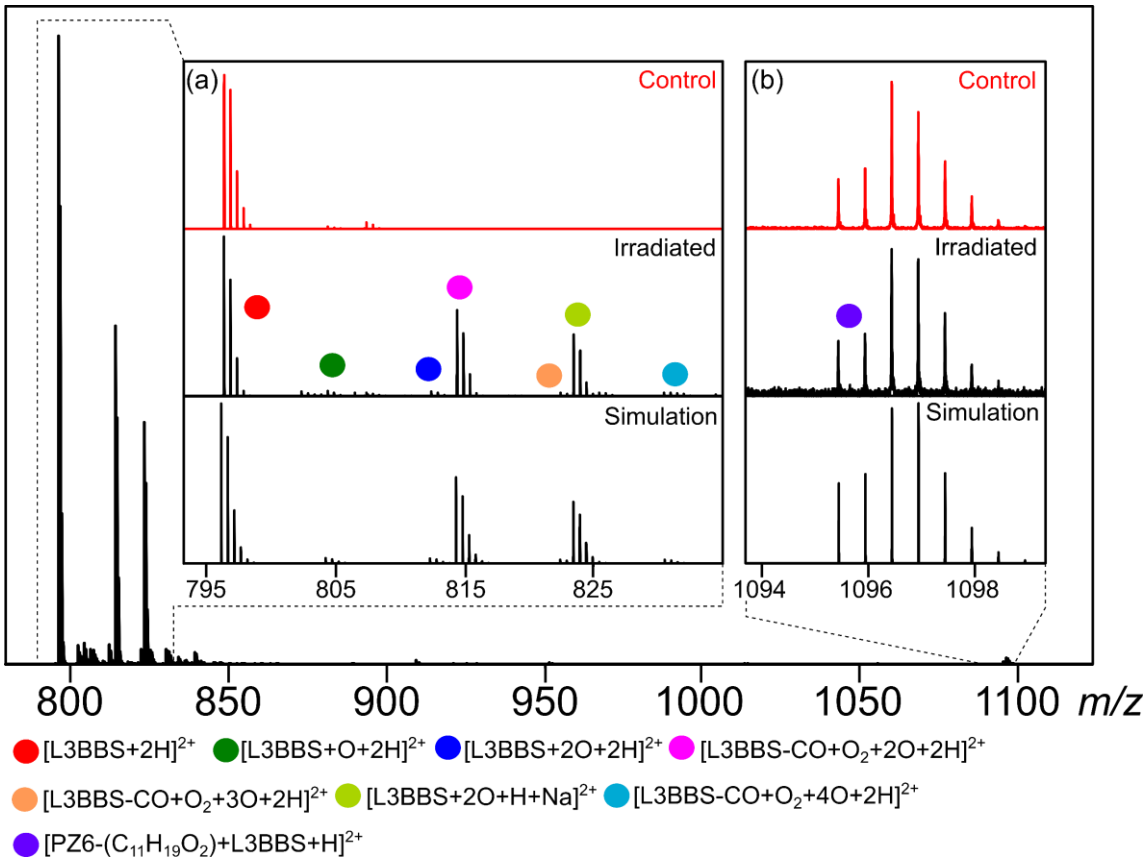


Figure 4.8 Full MS of PZ6-L3BBS (2:1) by nESI. Inset (a) shows the formation of oxidised products. Inset (b) shows the formation of iridium-bound products. Identities of oxidised products and iridium-bound products are listed below the figure.

Unlike PZ1, PZ3, PZ4, and PZ6, for PZ2 and PZ5 (Figures 4.9 and 4.10), which contained the same (C,N) ligand (shown in Figure 4.3), only oxidised products were formed and no iridium-bound product was observed, indicating variable products could be formed even when using complexes within the same synthetic series.

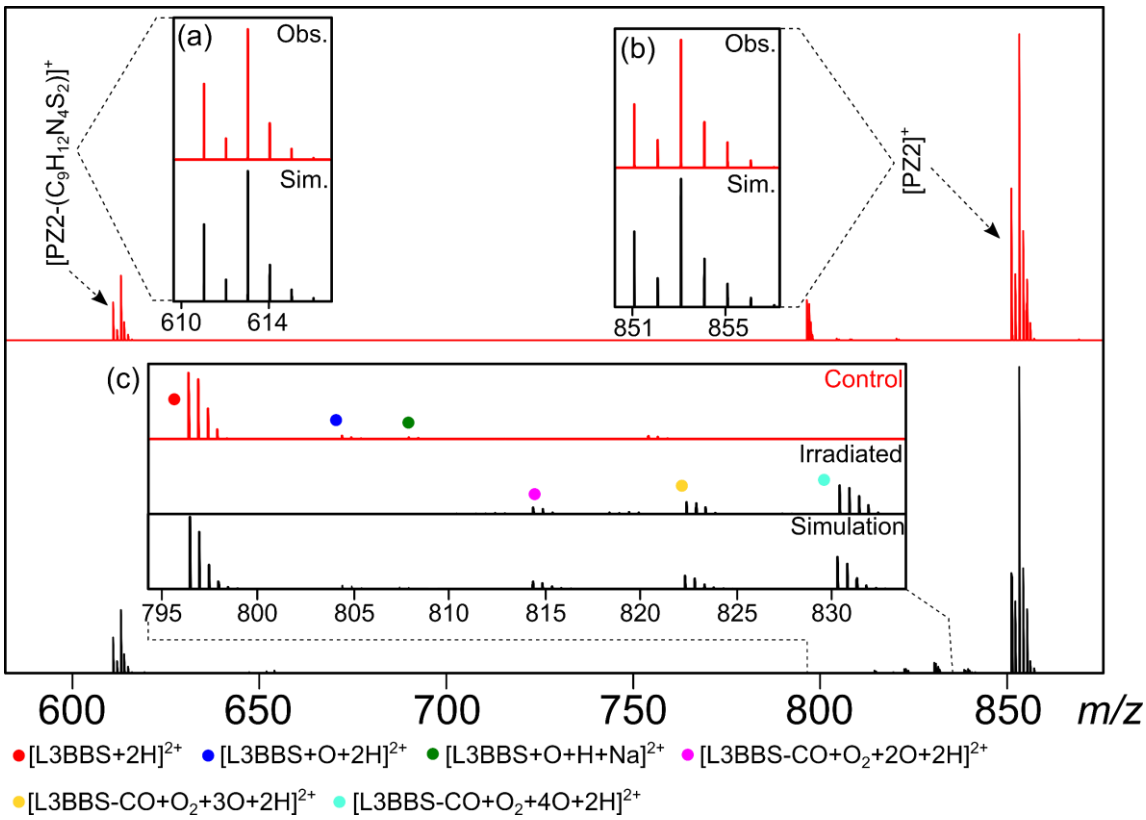


Figure 4.9 Full MS of PZ2-L3BBS (2:1) by nESI. Inset (a) shows the simulation of PZ2 with the loss of dithione ligand. Inset (b) shows simulation of intact PZ2. Inset (c) shows the formation of oxidised products. Identities of oxidised products and iridium-bound products are listed below the figure.

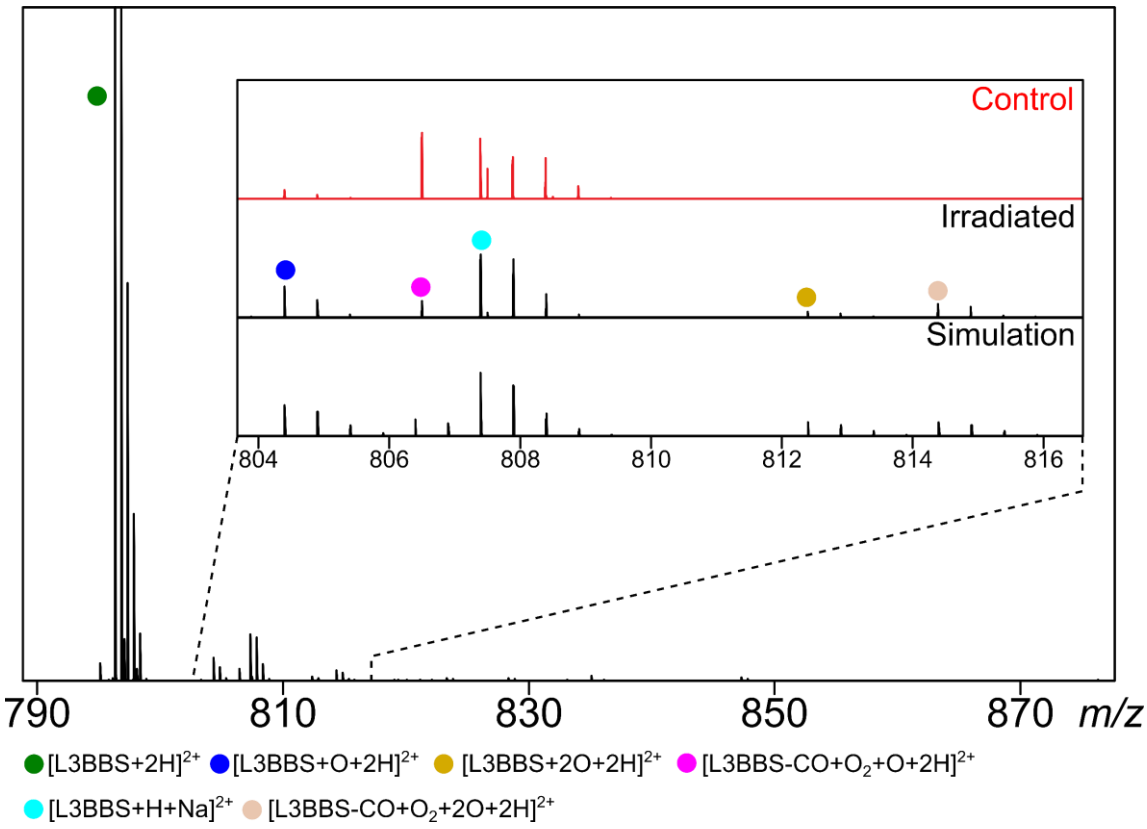


Figure 4.10 Full MS of PZ5-L3BBS (2:1) by nESI. Inset shows the formation of oxidised products. Identities of oxidised products and iridium-bound products are listed below the figure.

To further investigate the specific oxidation sites in the oxidised L3BBS products, a 5:1 reaction ratio of PZ6-L3BBS was prepared, to increase the amount of $^1O_2$ per L3BBS molecule in the reaction system, thus increasing the number of oxidised products. A full MS (Figure 4.11) was analysed to make sure there were no significant differences in products between the 2:1 and 5:1 reaction system. Up to 5 different oxidation products were observed. Apart from the oxidation products, PZ6 was also found with the loss of the diketonate ligand, and addition of one acetonitrile or two acetonitrile molecules were observed in the mass spectrum (Figure 4.11). To study the specific oxidation site on the oxidation products, tandem mass spectrometry (MS/MS) was required, to fragment the isolated oxidised species and identify binding sites.

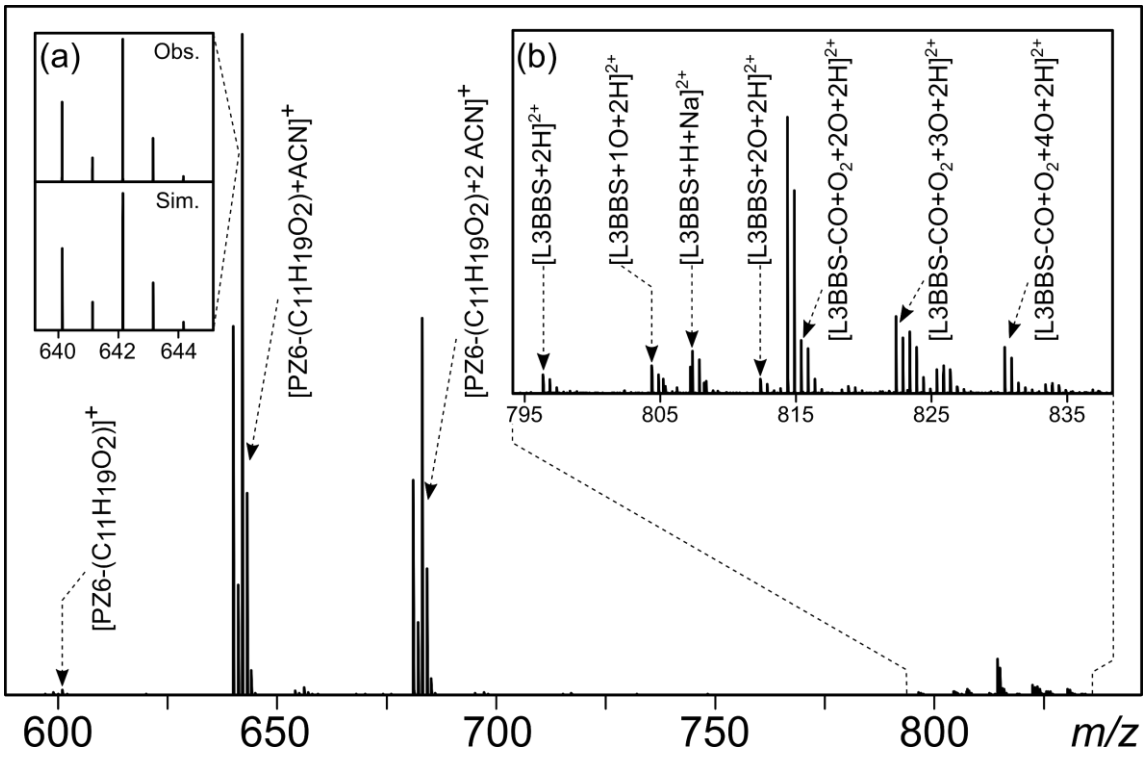


Figure 4.11 Full MS of PZ6-L3BBS (5:1) by nESI. Inset (a) shows the simulation of PZ6 with the loss of the diketonate ligand and an addition of an acetonitrile molecule. Inset (b) shows an array of doubly charged oxidised L3BBS products.

Specific oxidation sites were investigated by ECD MS/MS. Initially $[L3BBS+O+2H]^{2+}$, 804.409192 *m/z* was isolated in the quadrupole with an isolation window of 5 *m/z*, however, with the tight isolation window, the ion intensity decreased dramatically. Therefore, a wider isolation window of 10-20 *m/z* was used to isolate the ions of interest. Unfortunately, this led to another problem of failing to obtain a "clean" isolation of the ion of interest as the other oxidised products were close to each other within the target mass ranges. Therefore, an in-cell isolation technique, multi-CHEF was employed. The ion of interest (804.409192 *m/z*) was first isolated in the quadrupole with a wide isolation window of 30 *m/z*. Then the ions were transferred and accumulated in the hexapole-based collision cell, and later transferred into the infinity cell, where the second isolation event could be carried out. Ions of interest were carefully isolated again inside the infinity cell with a notch width of 0.5 *m/z* and with isolation power of 50%, which allowed to obtain a "clean" isolation of the ion of interest, and ECD was applied to fragment the precursor ion (Figure 4.12, Table S4.1).

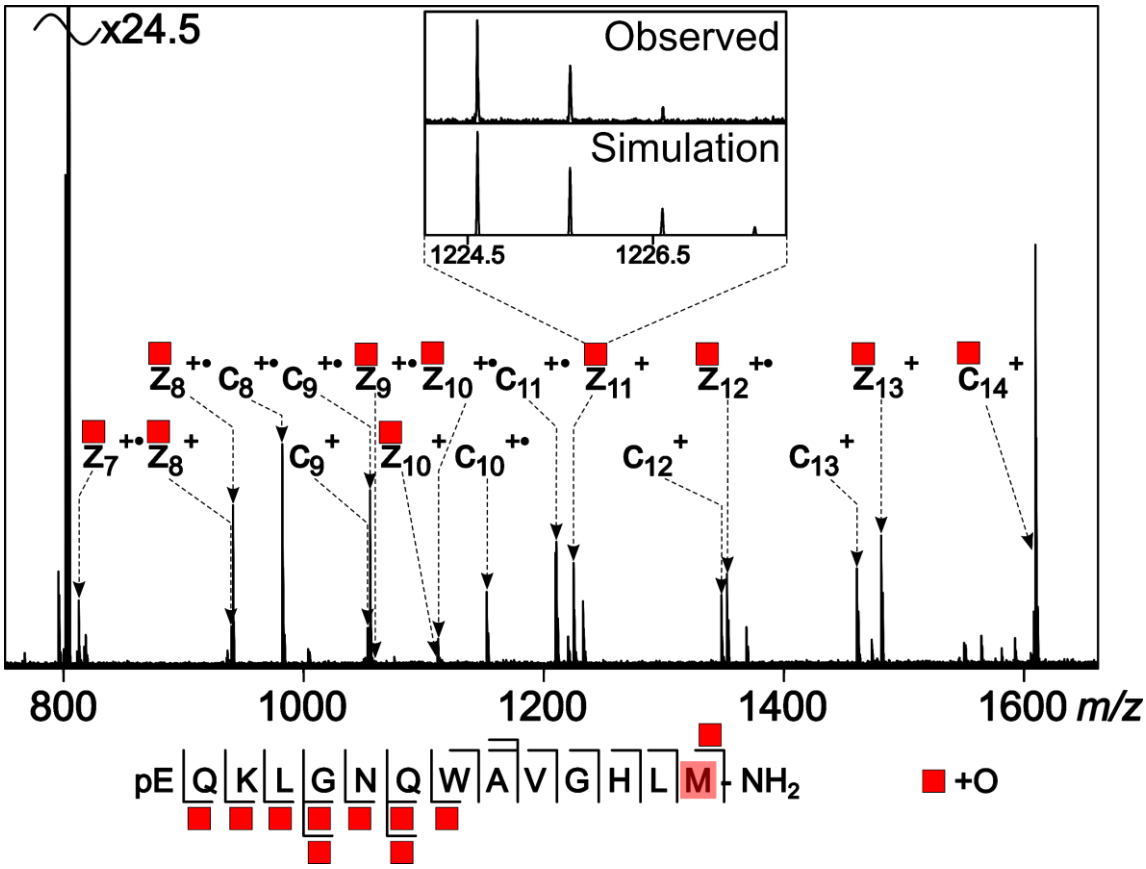


Figure 4.12 ECD MS/MS of $[L3BBS+O+2H]^{2+}$. Inset shows the simulation of the $[z_{11}+O]^{+}$ fragment ion. Red labels indicate the presence of oxidation. The oxidation site methionine is highlighted in red. The fragmentation map is shown below the MS/MS spectrum.

The mass difference between the unmodified L3BBS (796.411734 *m/*z) and the precursor (804.409192 *m/*z) is 15.994915 Da, which corresponds to an oxygen atom. However, there are multiple possible oxidation sites in the peptide sequence (pEQKLGNQWAVGHLM-NH_2), which are tryptophan, histidine, and methionine. With careful internal calibration and analysis of the MS/MS spectrum, peaks were assigned with 100% cleavage coverage. The fragments observed in the ECD MS/MS spectrum are summarised in the fragmentation map below the spectrum. Fragments labelled with a red square indicate the presence of oxidation. The oxidation site was located at the C-terminal methionine, indicating the methionine was oxidised once to methionine sulfoxide.

Subsequent oxidation products were then isolated in a similar fashion (combination of quadrupole isolation and in-cell isolation), and ECD was applied again to investigate the oxidation site (shown in Figure 13, Table S2).

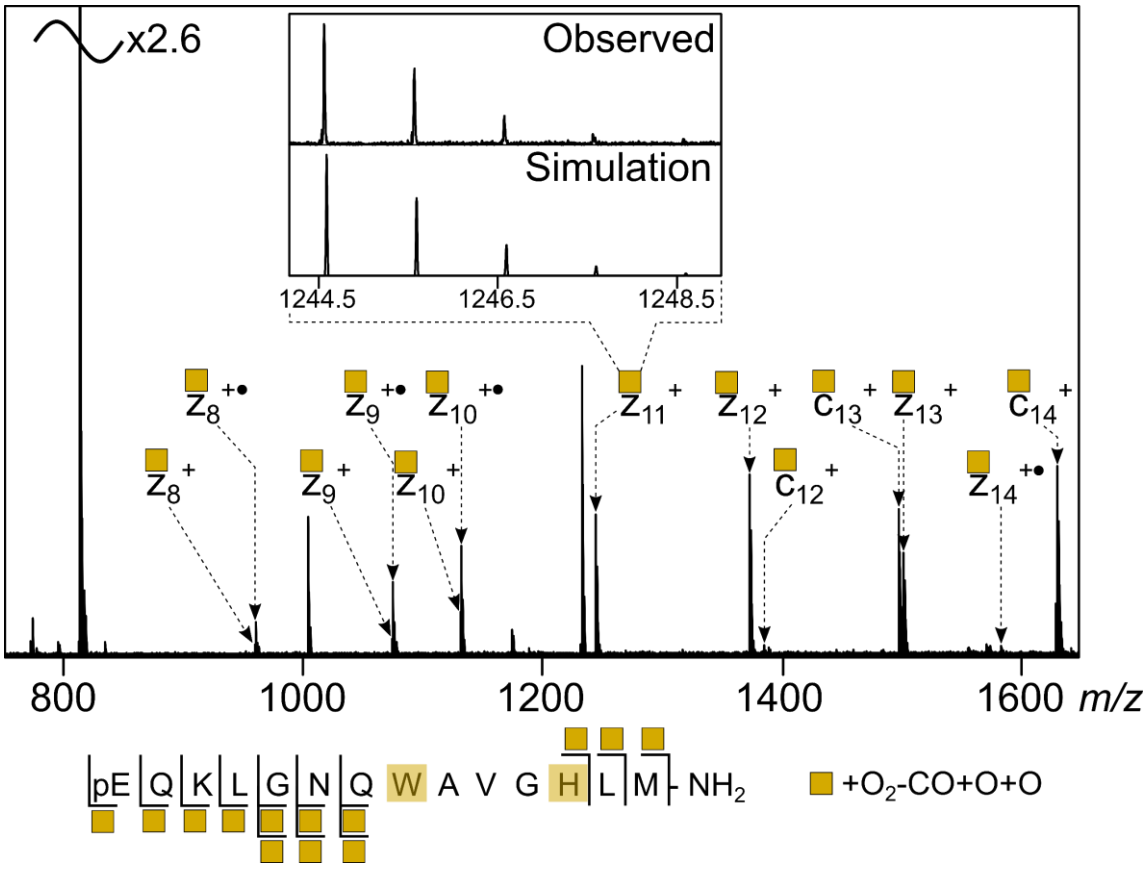


Figure 4.13 ECD MS/MS of $[L3BBS+4O-CO+2H]^{2+}$. Inset shows the simulation of the $[z_{11}+4O-CO]^{+}$ fragment ion. Yellow labels indicate the presence of oxidation. The oxidation sites tryptophan and histidine are highlighted in yellow. The fragmentation map is shown below the MS/MS spectrum.

A cleavage coverage of 62% was observed in the ECD MS/MS of the isolated oxidation product of $[L3BBS+4O-CO+2H]^{2+}$ (814.404106 *m/z*). The oxidation sites were located into the region containing residues QWAVGH. Although fragments obtained in the ECD MS/MS spectrum were not sufficient to narrow down the oxidation sites to individual residues, by analysing individual residue within the region, only tryptophan (W) and histidine (H) can be oxidised. However, histidine can only be oxidised once into 2-oxo-histidine, suggesting the remaining modification ($+O_2$-CO+O)

was on the tryptophan, thus, the formation of 3-hydroxykynurenine (3OH-kyn) (Figure 4.4). 2-Oxo-histidine is commonly observed when histidine is exposed to a singlet oxygen-rich environment and 3OH-kyn is a product induced by $^1O_2$ oxidation of tryptophan,[12,13,26,34] together provide a strong evidence of the generation of $^1O_2$ when the iridium-based photosensitisers are irradiated and activated with visible blue light (463 nm).

Two series of Ir(III)-based photosensitisers were designed to generate highly toxic $^1O_2$ upon irradiation with wavelength of 463 nm. $^1O_2$ generated in cells would greatly increase the extent of oxidation events and induce a vast amount of oxidative stress, potentially leading to apoptosis. These complexes were reacted with a model peptide, L3BBS, to investigate the MoA of these drugs towards biomolecules. A series of oxidation events and oxidised products were observed, which was suggested to be induced by $^1O_2$ by the formation of kynurenine/3-hydroxykynurenine from tryptophan and the observation of 2-oxo-histidine. Surprisingly, the iridium-complexes were found to react with L3BBS directly and form iridium-bonded L3BBS products, which could provide a bimodal strategy to kill cancer cells, lowering the probably of developing resistance against these photosensitisers.

PZ3 and PZ6 were shown to be more potent out of the six photosensitisers. Therefore, they were chosen to study in cells. PDT in cell experiments was first tested by incubating monolayer cancer cells with the photosensitisers (PZ3 and PZ6) at various concentrations for 2 h. In the dark, PZ6 showed no cytotoxicity to both A549 lung cancer and MRC-5 normal fibroblasts (IC_{50} > 100 μM), while PZ3 showed moderate toxicity to A549 cells (IC_{50} = 21.2 μM, Table S4.3). The potency of both PZ3 and PZ6 towards cancer cells increased markedly upon irradiation. Under the same experimental conditions, 5-aminolevulinic acid (5-ALA) and cisplatin displayed much lower phototoxicity (IC_{50} > 100 μM). The phototoxicities of PZ3 and PZ6 towards 3D multicellular tumour spheroids (MCTSs) were further investigated. Untreated MCTSs were unaffected by 1-photon (1P) (463 nm) and 2-photon (2P) (750 nm) irradiation. 1P refers to a linear absorption process where one photon excites an atom ion or molecule from a lower energy level to a higher energy state (excited state) and 2P is the

absorption of two photons of identical or different frequencies. PZ6 showed no toxicity towards A549 MCTSs in the dark, and no toxicity towards MRC-5 MCTSs in either the dark or upon irradiation. However, the IC_{50} values for PZ6 towards A549 spheroids upon 1P and 2P irradiation were very low (high potency), 1.0 μM and 0.23 μM, respectively (Table S4.3). Interestingly, 2P photocytotoxicity was higher than 1P photocytotoxicity for this compound. Although PZ3 also showed a decent phototoxicity to A549 spheroids, it also exhibited some phototoxicity towards the MRC-5 spheroids, i.e. insufficient selectivity. Therefore, PZ6 was chosen to use in the following proteomic study.

A549 lung cancer cell (ca. 1×10^9) were treated with PZ6 (10 μM, 2 h) either in the dark or with $\lambda = 463$ nm light irradiation. Proteins from these cells were harvested and digested as described in the experimental section. The digested proteins were analysed using a nLC-FT-ICR MS/MS setup equipped with a C18 reversed phase (RP) column. Each of the three samples (control, drug treated without irradiation, and drug treated with irradiation) were repeated five times for reproducibility and reliability purposes. The nLC-MS/MS data sets were searched against the human proteome database using the Mascot search engine, including oxidative modifications of the side chains of cysteine (C), histidine (H), tyrosine (Y), phenylalanine (F), methionine (M), and tryptophan (W), as well as asparagine (N) and glutamine (Q) deamidation and serine (S), threonine (T), and tyrosine (Y) phosphorylation. With 1% false discovery rate (FDR) against the database, all peptide identified by Mascot were manually checked for false positives. Only proteins which detected three times or more (in triplicates) out of the five technical replicates under all three conditions were quantified. 212 proteins and ~ 1500 peptides met these requirements and were identified (Figure 4.14).

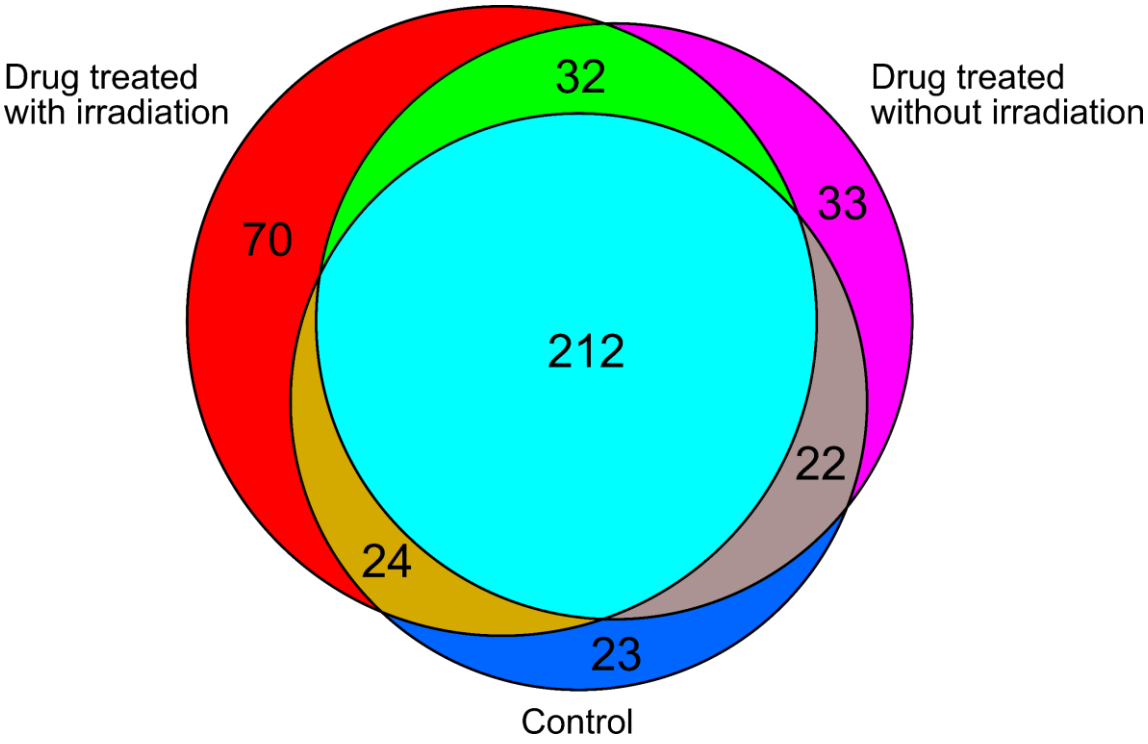


Figure 4.14 Venn diagram of proteins identified by Mascot under three different reaction conditions (control, drug treated without irradiation, and drug treated with irradiation).

Quantification was achieved using a label-free approach, by spiking a known concentration and volume of tryptic digested lysozyme as an internal control. A lysozyme peptide (FESNFNTQATNR, 714.8365 *m/z*) was used as a reference peptide to compare across different data sets. Two-tailed t-tests ensured the consistency of the reference peptide throughout all data sets, and all p values were calculated to be larger than 0.05 (i.e. quantities of the reference lysozyme peptide were statistically similar, and consistent throughout all data sets (Table S4.4).

Five unique oxidised peptides were detected three times or more out of the five replicates of nLC-MS/MS runs under all three reaction conditions. These peptides are observed from different proteins and quantified using a label-free approach. There are two commonly used label-free quantification methods in LC-MS, (a) Ion Abundance and (b) Ion Abundance Ratio (Figure 4.15). "Ion Abundance" quantification is achieved by comparing directly the peak areas of the extract ion chromatogram (EIC) of the ions

of interest (Figure 4.15a). However, this method is often limited by the consistency and reproducibility of the LC-MS runs, such as the ionisation efficiency of analytes due to the difference in emitter performance. "Ion Abundance Ratio" quantification refers to the comparison between the ratios obtained from the ion of interest and the internal control under different conditions (Figure 4.15b).

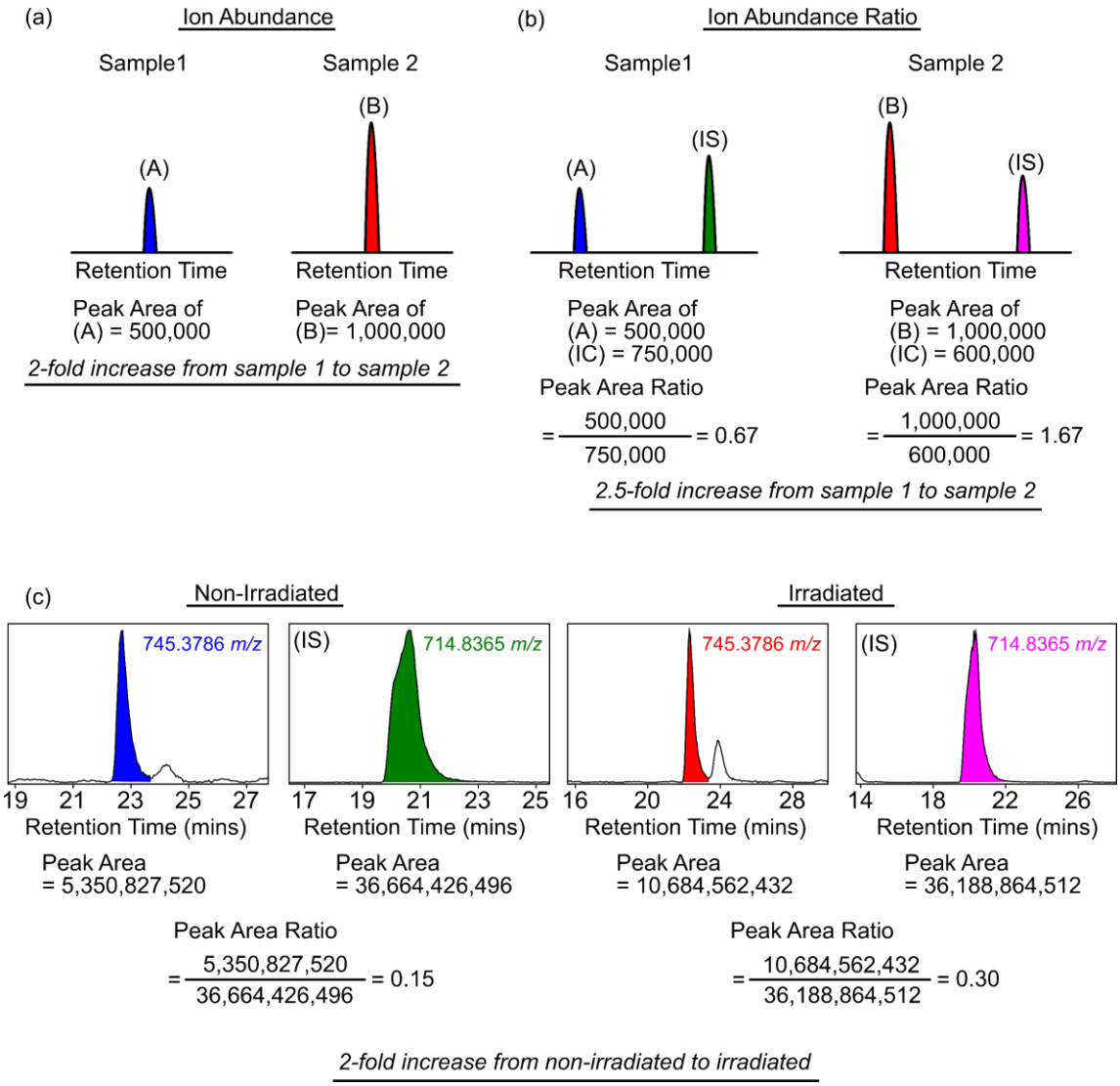


Figure 4.15 Label-free quantification. (a) Ion Abundance. (b) Ion Abundance Ratio. (c) Example of quantifying the oxidised aldose reductase peptide (YKPAVNQIECHPYLTQEK + O), 745.3786 *m/z*. IS represents internal standard.

$$Fold\ of\ Change = \frac{Average\ of\ the\ area\ ratios\ (irradiated)}{Average\ of\ the\ area\ ratios\ (dark)} \qquad \text{[Eq. 2]}$$

The five oxidised peptides were quantified using the "Ion Abundance Ratio" label-free quantification method with the lysozyme reference peptide (FESNFNTQATNR, 714.8365 *m/z*). T-tests were then conducted and only peptides with p values < 0.05 were considered to have significant change between the two different reaction conditions (drug treated without irradiation and drug treated with irradiation). Three out of five of the oxidised peptides found to have p values < 0.05, i.e. significantly different (Table S4.5). Two of the three unique peptides were upregulated. The level of oxidised peptide from heat shock protein 70 kDa (Hsp 70), Ala329-Arg442, AQIHDIVLVGGSTR (741.4156 *m/z*), increased by 5.66-fold (Table S4.5) for drug-treated cells upon irradiation compared to drug-treated cells in the dark. Specific oxidation of His332 to 2-oxo-His332 was identified by nLC-CAD MS/MS with cleavage coverage of 62% (Figure 4.16b, Table S4.6). The second oxidised peptide, Tyr178-Lys195, from aldose reductase, YKPAVNQIECHPYLTQEK (745.3780 *m/z*, with alkylated cysteine, Figure 4.17b, Table S4.7) contained 2-oxo-His188 (with 39% cleavage coverage using nLC-CAD MS/MS), which increased by 2.93-fold for drug-treated cells upon irradiation as compared to that in the dark (Table S4.5). These data appear to be the first report of the formation of 2-oxo-histidine after treatment of cancer cells with organometallic photosensitisers.[35]

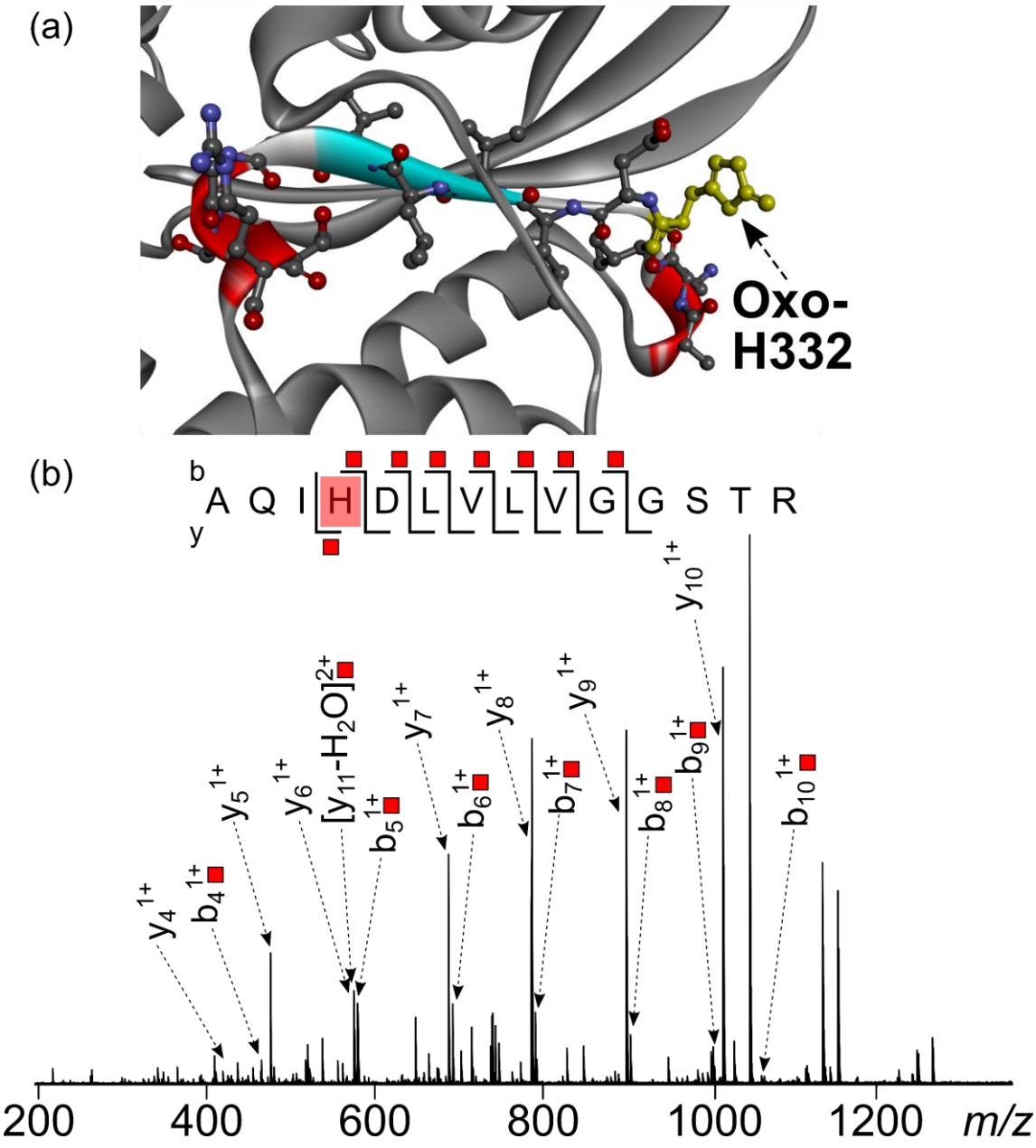


Figure 4.16 (a) Structure of Hsp 70 (PDB: 3ATV),[36] with the oxidised peptide Ala329-Arg342 shown in colour (2-oxo-His332, yellow). (b) nLC-CAD MS/MS of the oxidised peptide from Hsp 70. Fragments with red labels indicate the presence of oxidation. (Reproduced from Zhang et. al. 2017 with permission)[37].

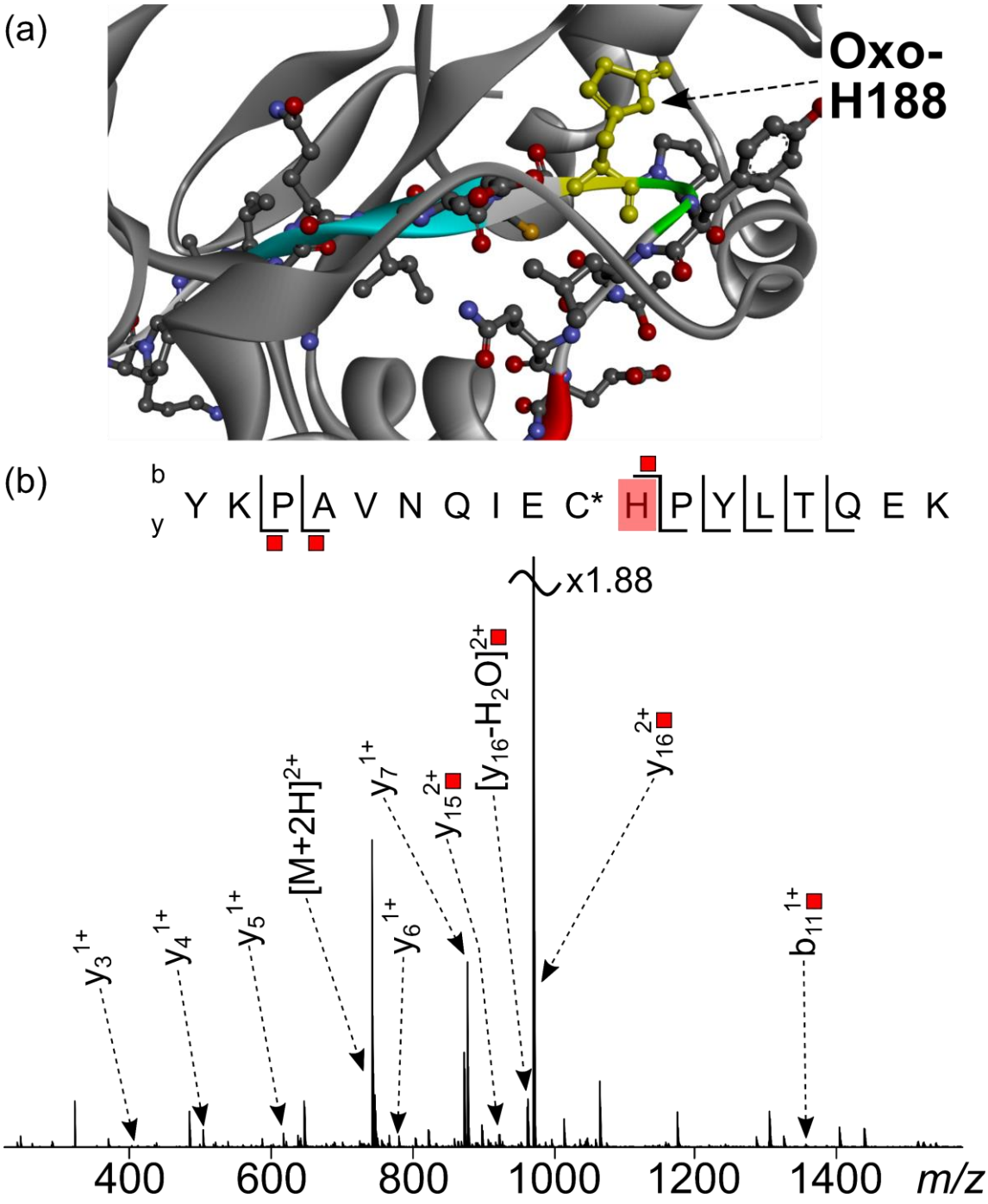


Figure 4.17 (a) Structure of aldose reductase (PDB: 1US0),[38] with the oxidised peptide Tyr178-Lys195 shown in colour (2-oxo-His188, yellow). (b) nLC-CAD MS/MS of the oxidised peptide from aldose reductase. Fragments with red labels indicate the presence of oxidation, and an asterisk indicates alkylated cysteine. (Reproduced from Zhang et. al. 2017 with permission)[37].

Pathway analysis was carried out to investigate the overall effect induced by PZ6 on cell metabolism (method described in the experimental section). The most significant result was the identification of nine unique proteins along the glycolysis pathway (Figure 4.18, Table S4.8). The levels of these proteins, which are all involved in the conversion of glucose to pyruvate, increased by factor of ~ 2.1-5.3-fold with irradiation of A549 cancer cells treated with PZ6, with the highest increased for fructose-biphosphate aldolase. Cancer cells have been shown to have defective mitochondria and increase their rate of glycolysis as a source of adenosine triphosphate (ATP) and energy to compensate for this mitochondrial effect.[39] Mitochondria, where oxygen is reduced to water, is also a major source of ROS in cells.[40] During irradiation, a vast amount of $^1O_2$ is generated and a loop of ROS-stimulated glucose uptake and glucose-stimulated ROS production is triggered.[41–43] This process is consistent with the up-regulation of proteins in the glycolytic pathway.

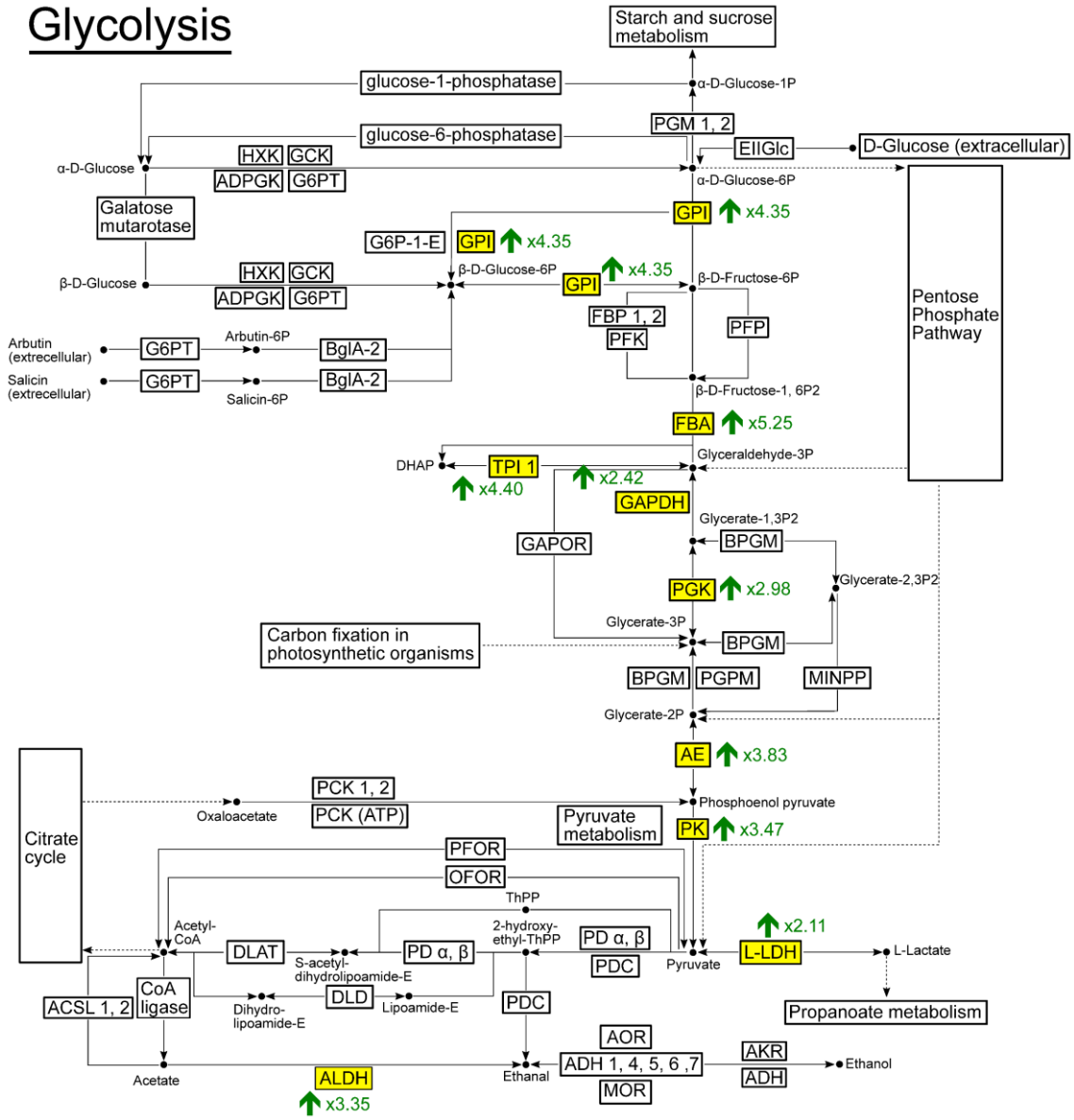


Figure 4.18 Effects of irradiation of PZ6 on the level of proteins associated with glycolysis pathway in A549 lung cancer cells. Nine up-regulated proteins highlighted in yellow were identified by nLC-CAD MS/MS, with the magnitude of the fold-change indicated in green. (Reproduced from Zhang et. al. 2017 with permission)[37].

Key: **PGM**: Phosphoglucomutase; **EIIGlc**: D-glucose Npi-phosphotransferase; **HXK**: Hexokinase; **GCK**: Glucokinase; **ADPGK**: ADP dependent glucokinase; **G6PT**: D-glucose-6-phosphotransferase; **G6P-1-E**: D-glucose-6-phosphate 1-epimerase; **GPI**: Glucose-6-phosphate isomerase; **BgIA-2**: 6-phospho-beta-D-glucosyl-D-glucose

glycohydrolase; **FBP**: Fructose-bisphosphatase; **PFK**: Phosphofructokinase; **PFP**: D-fructose-6-Phosphate 1-phosphotransferase; **FBA**: Fructose-bisphosphate aldolase; **TPI**: Triosephosphate isomerase; **GAPDH**: Glyceraldehyde-3-phosphate dehydrogenase; **GAPOR**: D-glyceraldehyde-3-phosphate: NADP+ oxidoreductase; **BPGM**: Bisphosphoglycerate mutase; **PGK**: Phosphoglycerate kinase; **PGPM**: D-phosphoglycerate 2,3-phosphomutase; **MINPP**: Multiple inositol-polyphosphate phosphatase; **AE**: Alpha-enolase; **PCK**: Phosphoenolpyruvate carboxykinase; **PCK (ATP)**: ATP Oxaloacetate carboxy-lyase; **PK**: pyruvate kinase; **PFOR**: Pyruvate ferredoxin-2-oxidoreductase; **OFOR**: 2-oxocarboxylate: ferredoxin-2-oxidoreductase; **ACSL**: Acyl-CoA-synthetase; **DLAT**: Dihydrolipoamide S-acetyltransferase; **DLD**: Dihydrolipoamide dehydrogenase; **PD**: Pyruvate dehydrogenase; **PDC**: 2-oxo-acid-carboxyl-lyase; **L-LDH**: Lactate dehydrogenase; **ALDH**: Aldehyde dehydrogenase; **ADH**: Alcohol dehydrogenase; **AOR**: Alcohol cytochrome c oxidoreductase; **MOR**: Methanol cytochrome c oxidoreductase; **AKR**: Aldo-keto reductase.

Conclusions

Six iridium-(III)-based photosensitisers were designed to generate singlet oxygen upon irradiation with specific wavelength (463 nm, visible blue light) and their mechanisms of reaction with the model peptide, [Lys^3]-Bombesin, were analysed by FT-ICR MS. Four out of the six iridium-based photosensitisers showed the ability to not only oxidise, but also chemically bind to the target peptide, which displayed the possibility of a dual mechanisms of action against target biomolecules, and lowering the chance of developing complete resistance against the photosensitisers. A series of oxidised products of the model peptide were observed and, by using ECD MS/MS, oxidation sites were found to be methionine, histidine, and tryptophan. A specific oxidised form of tryptophan, 3-hydroxykynurenine was identified and provided strong evidence of the generation of singlet oxygen, and the oxidised form of histidine, 2-oxo-histidine, was also often observed as a product upon singlet oxygen oxidation. PZ6 was designed to release the diketonate ligand upon irradiation and found to possess the highest selectivity among the six photosensitisers, i.e. little to no cytotoxicity in the dark and extremely potent when irradiated/activated.

In previous reports, the specific nature of the damage to proteins in the cell, induced by photodynamic therapy (PDT), has been little studied. Therefore, effect of PZ6 on A549 lung cancer cells was investigated by nLC-FT-ICR CAD MS/MS. Singlet oxygen generated by PZ6 was found to oxidise specific histidine residues in the proteins heat shock protein 70 kDa (Hsp 70) and aldose reductase (AR), which have important functions in cancer cells. Hsp 70 is a molecular chaperone for nascent proteins, and aberrantly folded, damaged, or mutated proteins. AR is a monomeric reduced nicotinamide adenine dinucleotide phosphate (NADPH)-dependent enzyme, a member of the aldo-keto reductase superfamily. This work appears to be the first report showing that specific sites of cellular Hsp 70 and AR can be oxidising during PDT. The combination of oxidative stress induced by photoactivation of PZ6 together with the malfunction of mitochondria in cancer cells leads to the increase use of glucose to generate energy and is consistent with the observation of increase in the levels of 9 unique enzymes involved in the glycolytic pathway (by a factor of ~ 2.1-5.3-fold).

References

(1) Sirichanchuen, B.; Pengsuparp, T.; Chanvorachote, P. Long-Term Cisplatin Exposure Impairs Autophagy and Causes Cisplatin Resistance in Human Lung Cancer Cells. *Mol. Cell. Biochem.* **2012**, *364* (1–2), 11–18.

(2) Zayed, A.; Shoeib, T.; Taylor, S. E.; Jones, G. D. D.; Thomas, A. L.; Wood, J. P.; Reid, H. J.; Sharp, B. L. Determination of Pt-DNA Adducts and the Sub-Cellular Distribution of Pt in Human Cancer Cell Lines and the Leukocytes of Cancer Patients, Following Mono- or Combination Treatments, by Inductively-Coupled Plasma Mass Spectrometry. *Int. J. Mass Spectrom.* **2011**, *307* (1–3), 70–78.

(3) Centerwall, C. R.; Tacka, K. A.; Kerwood, D. J.; Goodisman, J.; Toms, B. B.; Dubowy, R. L.; Dabrowiak, J. C. Modification and Uptake of a Cisplatin Carbonato Complex by Jurkat Cells. *Mol. Pharmacol.* **2006**, *70* (1), 348–355.

(4) Wootton, C. A.; Millett, A. J.; Lopez-Clavijo, A. F.; Chiu, C. K. C.; Barrow, M. P.; Clarkson, G. J.; Sadler, P. J.; O'Connor, P. B. Structural Analysis of Peptides Modified with Organo-Iridium Complexes, Opportunities from Multi-Mode Fragmentation. *Analyst* **2019**, *144*, 1575–1581.

(5) Novohradsky, V.; Zerzankova, L.; Stepankova, J.; Kisova, A.; Kostrhunova, H.; Liu, Z.; Sadler, P. J.; Kasparkova, J.; Brabec, V. A Dual-Targeting, Apoptosis-Inducing Organometallic Half-Sandwich Iridium Anticancer Complex. *Metallomics* **2014**, *6* (8), 1491–1501.

(6) Banerjee, S.; Soldevila-Barreda, J. J.; Wolny, J. A.; Wootton, C. A.; Habtemariam, A.; Romero-Canelon, I. New Activation Mechanism for Half-Sandwich Organometallic Anticancer Complexes. *Chem. Sci.* **2018**, *9*, 3177–3185.

(7) Wootton, C. A.; Sanchez-Cano, C.; Lopez-Clavijo, A. F.; Shaili, E.; Barrow, M. P.; Sadler, P. J.; O'Connor, P. B. Sequence-Dependent Attack on Peptides by Photoactivated Platinum Anticancer Complexes. *Chem. Sci.* **2018**, *9* (10), 2733–2739.

(8) Hearn, J. M.; Romero-Canelón, I.; Munro, A. F.; Fu, Y.; Pizarro, A. M.; Garnett, M. J.; McDermott, U.; Carragher, N. O.; Sadler, P. J. Potent Organo-Osmium Compound Shifts Metabolism in Epithelial Ovarian Cancer Cells. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112* (29), E3800–E3805.

(9) Wootton, C. A.; Sanchez-Cano, C.; Liu, H.-K.; Barrow, M. P.; Sadler, P. J.; O'Connor, P. B. Binding of an Organo-Osmium(II) Anticancer Complex to Guanine and Cytosine on DNA Revealed by Electron-Based Dissociations in High Resolution Top-Down FT-ICR Mass Spectrometry. *Dalton Trans.* **2015**, *44* (8), 3624–3632.

(10) Gabbiani, C.; Casini, A.; Kelter, G.; Cocco, F.; Cinellu, M. A.; Fiebig, H.-H.; Messori, L. Mechanistic Studies on Two Dinuclear Organogold(III) Compounds Showing Appreciable Antiproliferative Properties and a High Redox Stability. *Metallomics* **2011**, *3* (12), 1318–1323.

(11) Wills, R. H.; Habtemariam, A.; Lopez-Clavijo, A. F.; Barrow, M. P.; Sadler, P. J.; O'Connor, P. B. Insights into the Binding Sites of Organometallic Ruthenium Anticancer Compounds on Peptides Using Ultra-High Resolution Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2014**, *25* (4), 662–672.

(12) Agon, V. V.; Bubb, W. A.; Wright, A.; Hawkins, C. L.; Davies, M. J. Sensitizer-Mediated Photooxidation of Histidine Residues: Evidence for the Formation of Reactive Side-Chain Peroxides. *Free Radic. Biol. Med.* **2006**, *40* (4), 698–710.

(13) Oattison, D. I.; Rahmanto, A. S.; Davis, M. J. Photo-Oxidation of Proteins. *Photochem. Photobiol. Sci.* **2012**, *11*, 38.

(14) Nam, J. S.; Kang, M. G.; Kang, J.; Park, S. Y.; Lee, S. J. C.; Kim, H. T.; Seo, J. K.; Kwon, O. H.; Lim, M. H.; Rhee, H. W.; et al. Endoplasmic Reticulum-Localized Iridium(III) Complexes as Efficient Photodynamic Therapy Agents *via* Protein Modifications. *J. Am. Chem. Soc.* **2016**, *138* (34), 10968–10977.

(15) Farrer, N. J.; Salassa, L.; Sadler, P. J. Photoactivated Chemotherapy (PACT): The Potential of Excited-State d-Block Metals in Medicine. *Dalton Trans.* **2009**, No.

48, 10690–10701.

(16) Handing, K. B.; Niedzialkowska, E.; Shabalin, I. G.; Kuhn, M. L.; Zheng, H.; Minor, W. Characterizing Metal-Binding Sites in Proteins with X-Ray Crystallography. *Nat. Protoc.* **2018**, *13* (5), 1062–1090.

(17) Vinje, J.; Sletten, E. NMR Spectroscopy of Anticancer Platinum Drugs. *Anticancer. Agents Med. Chem.* **2007**, *7*, 35–54.

(18) Wilm, M. Principles of Electrospray Ionization. *Mol. Cell. Proteomics* **2011**, *10* (7), M111.009407.

(19) Wilm, M.; Mann, M. Analytical Properties of the Nanoelectrospray Ion Source. *Anal. Chem.* **1996**, *68* (1), 1–8.

(20) Jennings, K. R. Colllision-Induced Decompositions of Aromatic Ions. *Int. J. Mass Spectrom. Ion Phys.* **1968**, *1* (3), 227–235.

(21) Zubarev, R.; Kelleher, N. L.; McLafferty, F. W. Electron Capture Dissociation of Multiply Charged Protein Cations. A Nonergodic Process. *J. Am. Chem. Soc* **1998**, *120* (16), 3265–3266.

(22) Little, D. P.; Speir, J. P.; Senko, M. W.; O'Connor, P. B.; McLafferty, F. W. Infrared Multiphoton Dissociation of Large Multiply Charged Ions for Biomolecule Sequencing. *Anal. Chem.* **1994**, *66* (18), 2809–2815.

(23) Greer, S. M.; Brodbelt, J. S. Top-Down Characterization of Heavily Modified Histones Using 193 Nm Ultraviolet Photodissociation Mass Spectrometry. *J. Proteome Res.* **2018**, *17* (3), 1138–1145.

(24) Chen, Y.; Qiao, L.; Ji, L.; Chao, H. Phosphorescent Iridium(III) Complexes as Multicolor Probes for Specific Mitochondrial Imaging and Tracking. *Biomaterials* **2014**, *35* (1), 2–13.

(25) de Koning, L. J.; Nibbering, N. M. M.; van Orden, S. L.; Laukien, F. H. Mass Selection of Ions in a Fourier Transform Ion Cyclotron Resonance Trap Using Correlated Harmonic Excitation Fields (CHEF). *Int. J. Mass Spectrom.* **1997**,

165, 209–219.

(26) Davies, M. J. Singlet Oxygen-Mediated Damage to Proteins and Its Consequences. *Biochem. Biophys. Res. Commun.* **2003**, *305* (3), 761–770.

(27) Pattison, D. I.; Rahmanto, A. S.; Davies, M. J. Photo-Oxidation of Proteins. *Photochem. Photobiol. Sci.* **2012**, *11* (1), 38–53.

(28) Huang, D. W.; Sherman, B. T.; Lempicki, R. A. Systematic and Integrative Analysis of Large Gene Lists Using DAVID Bioinformatics Resources. *Nat. Protoc.* **2009**, *4* (1), 44–57.

(29) Kanehisa, M.; Goto, S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* **2000**, *2844* (1), 27–30.

(30) Kanehisa, M.; Sato, Y.; Kawashima, M.; Furumichi, M.; Tanabe, M. KEGG as a Reference Resource for Gene and Protein Annotation. *Nucleic Acids Res.* **2016**, *44* (D1), D457–D462.

(31) Kanehisa, M.; Furumichi, M.; Tanabe, M.; Sato, Y.; Morishima, K. KEGG: New Perspectives on Genomes, Pathways, Diseases and Drugs. *Nucleic Acids Res.* **2017**, *45* (D1), D353–D361.

(32) Davies, M. J. The Oxidative Environment and Protein Damage. *Biochim. Biophys. Acta - Proteins Proteomics* **2005**, *1703* (2), 93–109.

(33) B. Ohlsson, N. Fredäng, J. A. The Effect of Bombesin, Cholecystokinin, Gastrin, and Their Antagonists on Proliferation of Pancreatic Cancer Cell Lines. *Scand. J. Gastroenterol.* **2002**, *34* (12), 1224–1229.

(34) Gracanin, M.; Hawkins, C. L.; Pattison, D. I.; Davies, M. J. Singlet-Oxygen-Mediated Amino Acid and Protein Oxidation: Formation of Tryptophan Peroxides and Decomposition Products. *Free Radic. Biol. Med.* **2009**, *47* (1), 92–102.

(35) Uchida, K.; Kawakishi, S. 2-Oxo-Histidine as a Novel Biological Marker for Oxidatively Modified Proteins. *FEBS Lett.* **1993**, *332* (3), 208–210.

(36) Arakawa, A.; Handa, N.; Shirouzu, M.; Yokoyama, S. Biochemical and Structural Studies on the High Affinity of Hsp70 for ADP. *Protein Sci.* **2011**, *20* (8), 1367–1379.

(37) Zhang, P.; Chiu, C. K. C.; Huang, H.; Lam, Y. P. Y.; Habtemariam, A.; Malcomson, T.; Paterson, M. J.; Clarkson, G. J.; O'Connor, P. B.; Chao, H.; et al. Organoiridium Photosensitizers Induce Specific Oxidative Attack on Proteins within Cancer Cells. *Angew. Chemie Int. Ed.* **2017**, 14898–14902.

(38) Howard, E. I.; Sanishvili, R.; Cachau, R. E.; Mitschler, A.; Chevrier, B.; Barth, P.; Lamour, V.; Van Zandt, M.; Sibley, E.; Bon, C.; et al. Ultrahigh Resolution Drug Design I: Details of Interactions in Human Aldose Reductase-Inhibitor Complex at 0.66 Å. *Proteins Struct. Funct. Genet.* **2004**, *55* (4), 792–804.

(39) Wallace, D. C. Mitochondria and Cancer. *Nat Rev, Cancer* **2012**, *12*, 685–698.

(40) Jungwirth, U.; Kowol, C. R.; Keppler, B. K.; Hartinger, C. G.; Berger, W.; Heffeter, P. Anticancer Activity of Metal Complexes: Involvement of Redox Processes. *Antioxid. Redox Signal.* **2011**, *15* (4), 1085–1127.

(41) Liemburg-Apers, D. C.; Willems, P. H. G. M.; Koopman, W. J. H.; Grefte, S. Interactions between Mitochondrial Reactive Oxygen Species and Cellular Glucose Metabolism. *Arch. Toxicol.* **2015**, *89* (8), 1209–1226.

(42) Ghesquière, B.; Wong, B. W.; Kuchnio, A.; Carmeliet, P. Metabolism of Stromal and Immune Cells in Health and Disease. *Nature* **2014**, *511* (7508), 167–176.

(43) Zhang, C.; Cao, S.; Toole, B. P.; Xu, Y. Cancer May Be a Pathway to Cell Survival under Persistent Hypoxia and Elevated ROS: A Model for Solid-Cancer Initiation and Early Development. *Int. J. Cancer* **2015**, *136* (9), 2001–2011.

Supplementary Information

Table S4.1 Assignments for Figure 4.12: ECD MS/MS of $[L3BBS+O+2H]^{2+}$.

Oxidised species are highlighted in red.

Assignment	Charge State	Theoretical Mass (m/z)	Observed Mass (m/z)	Mass Error (ppm)
[z+1]7	1+	813.407606	813.40760	-0.01
z8	1+	940.458359	940.45845	0.10
[z+1]8	1+	941.466184	941.46622	0.04
[c-1]8	1+	998.492830	998.49303	0.20
[c-1]9	1+	1053.535029	1053.53523	0.19
c9	1+	1054.542854	1054.54280	-0.05
[z+1]9	1+	1055.509111	1055.50939	0.26
z10	1+	1111.522750	1111.52186	-0.80
[z+1]10	1+	1112.530575	1112.53016	-0.37
[c-1]10	1+	1152.603443	1152.60251	-0.81
[c-1]11	1+	1209.624906	1209.62494	0.03
z11	1+	1224.606814	1224.60734	0.43
c12	1+	1347.691643	1347.69148	-0.12
[z+1]12	1+	1353.709602	1353.71058	0.72
c13	1+	1460.775707	1460.77509	-0.42
z13	1+	1464.765440	1464.76623	0.54
c14	1+	1607.811107	1607.81110	0.00
			Average	0.00 ppm
			Absolute Average	0.30 ppm
			Standard Deviation	0.42 ppm

Table S4.2 Assignments for Figure 4.13: ECD MS/MS of $[L3BBS+4O-CO+2H]^{2+}$. Oxidised species are highlighted in red.

Assignment	Charge State	Theoretical Mass (m/z)	Observed Mass (m/z)	Mass Error (ppm)
z8	1+	960.448188	960.44815	-0.04
[z+1]8	1+	961.456013	961.45652	0.53
z9	1+	1074.491115	1074.49100	-0.11
[z+1]9	1+	1075.498940	1075.49935	0.38
z10	1+	1131.512579	1131.51286	0.25
[z+1]10	1+	1132.520404	1132.52059	0.16
z11	1+	1244.596643	1244.59628	-0.29
z12	1+	1372.691606	1372.69140	-0.15
c12	1+	1383.676387	1383.67658	0.14
c13	1+	1496.760451	1496.76071	0.17
z13	1+	1516.768908	1516.76887	-0.03
[z+1]14	1+	1613.797862	1613.79652	-0.83
c14	1+	1627.800936	1627.80198	0.64
			Average	0.06 ppm
			Absolute Average	0.29 ppm
			Standard Deviation	0.38 ppm

Table S4.3 Cytotoxicity (IC_{50}, μM) of the compounds towards 2D and 3D human A549 lung cancer and healthy MRC-5 lung cells.

	IC50 (μM)									
	2D A549		2D MRC-5		3D A549 Spheroids			3D MRC-5 Spheroids		
Compound	Dark	1P (463 nm)	Dark	1P (463 nm)	Dark	1P (463 nm)	2P (750 nm)	Dark	1P (463 nm)	2P (750 nm)
PZ3	21.2 ± 1.2	0.12 ± 0.04	48.9 ± 5.6	44.5 ± 2.4	30.8 ± 1.4	0.8 ± 0.1	0.20 ± 0.03	54.6 ± 0.9	46.9 ± 3.5	40.7 ± 3.6
PZ6	> 100	0.3 ± 0.1	> 100	> 100	> 100	1.0 ± 0.2	0.23 ± 0.10	> 100	> 100	> 100
5-ALA	> 100	> 100	> 100	> 100	> 100	> 100	N.D.	> 100	> 100	N.D.
Cisplatin	> 50	> 50	> 50	> 50	> 50	> 50	N.D.	> 50	> 50	N.D.

2D cells or 3D spheroids incubated with the compounds for 2 hours, medium replaced in both "dark" and "light" plates with fresh non-drug medium, "light" plates irradiated with blue 96-array LEDs (463 nm, 2.88 J/cm^2) or 2-photon laser beam (750 nm, 10 J/cm^2, 100 fs). All plates incubated for another 46 hours. N.D. = Not Determinable.

Table S4.4 T-test results for the lysozyme reference peptide (FESNFNTQATNR, 714.8365 *m/z*).

T Test bewteen	p value
Control and Drug treated without Irradiation	0.42
Control and Drug treated with Irradiation	0.16
Drug treated without irradiation and Drug Treated with irradiation	0.26

Table S4.5 Identification and changes in cellular levels of oxidised peptides detected by nLC-MS/MS in 5 triplicate samples, drug-treated without irradiation versus drug-treated irradiated. Peptides labelled in green are up-regulated, in red down-regulated, and black indicates peptides with p values > 0.05.

Protein Entry	Protein	Peptide Sequence	*m/z*	p value	Fold of change	Modification
A0A0G2JIW1	Heat Shock Protein 70kDa	AQIHDIVLVGGSTR	741.4156	0.0029	5.66 x [a]	Oxidation (H)
E9PCX2	Aldose reductase	YKPAVNQIECHPYLTQEK	745.3786	0.0013	2.93 x [b]	Oxidation (H)
A6NL76	Actin, muscle protein	QEYDEAGPSIVHR	758.8643	0.016	0.34 x	Oxidation (H)
Q9BQE3	Tubulin alpha chain	RAFVHWYVGEGMEEGEFSEAR	845.3690	0.073	0.64 x	Oxidation (H, W, M)
P04350	Tubulin beta chain	GHYTEGAELVDAVLDVVR	653.6729	0.46	1.20 x	Oxidation (H)

[a] Fold change = Area ratio (irradiated)/Area ratio (dark) = 0.293/0.0518 = 5.66 x

[b] Fold change = Area ratio (irradiated)/Area ratio (dark) = 0.450/0.154 = 2.93 x

Table S4.6 Assignments for Figure 4.16: nLC-CAD MS/MS of heat shock protein 70 kDa oxidised peptide, AQIHDIVLVGGSTR. Fragments with oxidation modification are labelled in red.

Assignment	Charge State	Theoretical Mass (m/z)	Observed Mass (m/z)	Mass Error (ppm)
[y4]	1+	420.220123	420.22018	0.14
[b4+O]	1+	466.240858	466.24092	0.13
[y5]	1+	477.241586	477.24147	-0.24
[y6]	1+	576.310000	576.31012	0.21
[y11+O-H_2O]	2+	576.312012	576.31199	-0.04
[b5+O]	1+	581.267801	581.26783	0.05
[y7]	1+	689.394064	689.39433	0.39
[b6+O]	1+	694.351865	694.35161	-0.37
[y8]	1+	788.462478	788.46226	-0.28
[b7+O]	1+	793.420279	793.41971	-0.72
[y9]	1+	901.546542	901.54651	-0.04
[b8+O]	1+	906.504343	906.50455	0.23
[b9+O]	1+	1005.572757	1005.57262	-0.14
[y10]	1+	1016.573485	1016.57398	0.49
[b10+O]	1+	1062.594221	1062.59413	-0.09
			Average	-0.02 ppm
			Absolute Average	0.24 ppm
			Standard Deviation	0.31 ppm

Table S4.7 Assignments for Figure 4.17: nLC-CAD MS/MS of aldose reductase oxidised peptide, YKPAVNQIECHPYLTQEK. Fragments with oxidation modification are labelled in red.

Assignment	Charge State	Theoretical Mass (m/z)	Observed Mass (m/z)	Mass Error (ppm)
[y3]	1+	404.213975	404.21391	-0.16
[y4]	1+	505.261653	505.26167	0.03
[y5]	1+	618.345717	618.34585	0.22
[M+3H]	3+	745.368528	745.36849	-0.05
[y6]	1+	781.409046	781.40880	-0.31
[y7]	1+	878.461810	878.46207	0.30
[y15+O]	2+	923.443625	923.44316	-0.50
[y16+O-H_2O]	2+	962.964725	962.96462	-0.11
[y16+O]	2+	971.970007	971.97022	0.22
[b11+O]	2+	1356.636497	1356.63611	-0.29
			Average	-0.07 ppm
			Absolute Average	0.22 ppm
			Standard Deviation	0.26 ppm

Table S4.8 Quantification of proteins in the glycolytic pathway, the level of 22 peptides from 9 proteins increase significantly as a result of 465 nm blue visible light irradiation of A549 human lung cancer cells treated with PZ6.

Protein [a]	Peptide	p value	Fold of Change	Average Change [b]
ALDH	IQQLEALQR	5.01E-03	+3.35	+3.35
AE	LNVTEQEKIDK	7.45E-03	+3.83	+3.83
FBA	GILAADESTGSIAKR	9.20E-04	+5.25	+5.25
GPI	HFVALSTNTTK	9.79E-05	+4.35	+4.35
GAPDH	VVDLMAHMASKE	3.01E-02	+2.42	+2.42
L-LDH	IVADKDYSVTANSK	8.19E-03	+1.97	+2.11
	VIGSGCNLDSAR	1.30E-02	+2.00	
	VTLTSEEEAR	4.29E-04	+2.81	
	DLADELALVDVIEDKLK	1.80E-02	+2.36	
PGK	AHSSMVGVNLPQK	3.59E-03	+2.98	+2.98
PK	AGKPVICATQMLESMIK	5.02E-03	+3.79	+3.47
	FGVEQDVDMVFASFIR	2.01E-02	+3.56	
	GSGTAEVELKK	3.83E-02	+2.04	
	GVNLPGAAVDLPAVSEKDIQDLK	9.88E-03	+10.59	
	ITLDNAYMEK	9.18E-03	+3.96	
	LNFSHGTHEYHAETIK	8.36E-04	+3.60	
	RFDEILEASDGIMVAR	3.01E-02	+2.02	
	TATESFASDPILYRPVAVALDTK	3.31E-02	+2.11	
TPI	IAVAAQNCYK	1.46E-03	+6.28	+4.40
	IIYGGSVTGATCK	1.14E-02	+1.79	
	SNVSDAVAQSTR	2.20E-04	+4.82	
	VVLAYEPVWAIGTGK	2.36E-03	+6.49	

[a] ALDH: Aldehyde dehydrogenase; AE: Alpha-enolase; FBA: Fructose-bisphosphate aldolase; GPI: Glucose-6-phosphate isomerase; GAPDH: Glyceraldehyde-3-phosphate dehydrogenase; L-LDH: L-lactate dehydrogenase; PGK: Phosphoglycerate kinase; PK: Pyruvate kinase; TPI: Triosephosphate isomerase. [b] Weighted average.

Chapter 5

Metallodrug-Proteomics Pathway Analysis using nLC-FT-ICR MS

Quantitative nLC-FT-ICR MS/MS was used to study the proteomic effects of an osmium-based anticancer drug against A2780 ovarian cancer cells and an iridium-based antimicrobial complex against *Staphylococcus aureus*. Promising results are reported in this chapter.

The piano-stool osmium-based anticancer complex, $[(\eta^6\text{-bip})Os(azpy\text{-}NMe_2)I]PF_6$, attacks energy production when administered to A2780 human ovarian cancer cells. Pyruvate was selectively accumulated, allowing increased energy production for extra amino acid biosynthesis and cell division.

An iridium-based antimicrobial complex, $[(Cp^{Xbiph})Ir(TolBig)Cl]Cl$, is shown to trigger a series of secondary responses from the bacterial cells by inducing a vast amount of stress, including oxidative stress. Extra protein/DNA biosynthesis events were triggered by the observation of up-regulated ribosomal proteins. These responses require a large amount of energy which was supported by the up-regulations of energy production-related enzymes.

The osmium-(II)-based anticancer complex was synthesised by Dr. Russell Needham and the drug treatment of A2780 ovarian cancer cell culture was carried out by Hannah Bridgewater. The iridium-(III)-based antimicrobial complex was synthesised and the drug treatment of *S. aureus* cell culture was carried out by Dr. Feng Chen. Sample preparation and all data were produced, analysed, and written by the thesis Author.

Introduction

Drug resistance is a critical problem in medical treatments which results in lowered effectiveness of drugs and more importantly, lowers the survival rate in patients.[1] Metallodrugs are commonly used against different diseases, such as bacterial infection[2] and different cancers.[3,4] These metallodrugs are usually transition metal-based, *e.g.* iridium,[5,6] osmium,[3,7] ruthenium,[8,9] gold,[10,11] rhodium,[12,13] platinum.[14,15] With the advantage of possessing multiple oxidation states, these metallodrugs are often designed as pro-drugs, and activated with specific stimulations, such as change in chemical environment (temperature, pH, salt concentration), or even with light (often used in photoactivatable chemotherapy or photodynamic therapy),[6,16,17] to enhance the selectivity towards the desired targets, such as bacterial or cancerous cells. These metallodrugs often possess different mechanisms of action (MoA) to overcome drug resistance and improve the effectiveness and increase the survival rate in patients. The study of new MoA's for these metallodrugs is often achieved by monitoring the changes in cells, such as the protein content of cells. The change in protein content can often provide insights into which metabolic pathways the metallodrugs affect in cells.

Proteomics is the study of the proteome and mass spectrometry (MS) is one the most commonly used analytical techniques because of its capability to analyse complex mixtures of analytes. Intact protein analysis (top-down proteomics) requires mass spectrometers with very high performance which are often very expensive, and therefore, bottom-up analysis is more commonly used. Bottom-up proteomics refers to the study of enzymatically digested proteins. Bottom-up proteomics creates even more species prior to MS analysis compared to intact protein analysis because each protein is digested into dozens of smaller peptides and often require more clean-up steps but is still the most employed approach to proteomics as the performance requirement of instruments is the lowest. Due to the complexity of proteomics samples, online or offline separation steps such as liquid chromatography (LC) are often used to separate the sample mixtures prior to MS analysis, in order to "simplify" the complexity of the protein/peptide mixtures, making the analysis more feasible. Reversed phase (RP) chromatography is the most common separation method used in LC proteomics studies

based on the hydrophobicity of the proteins/peptides. The more hydrophobic a species is, the longer the species is retained in the column and eluted at later time compared to hydrophilic species. The eluted species are then ionised, transmitted, and detected by the MS detector. Tandem mass spectrometry (MS/MS) allows fragmentation of the analyte species and can identify them. Different MS/MS fragmentation techniques, such as collisionally activated dissociation (CAD),[18] electron capture dissociation (ECD),[19] infrared multiphoton dissociation (IRMPD),[20] ultraviolet photodissociation (UVPD)[21], have been developed to provide informative fragments. Fourier transform-ion cyclotron resonance mass spectrometry (FT-ICR MS) not only provides the highest resolution and mass accuracy, but also possesses the largest array of fragmentation techniques.

In addition to protein/peptide identification, quantification is one of the main focuses of proteomics studies. Label-free quantification allows quantification of the protein/peptide content without the need of introducing chemical tags into the sample. Assumptions are made that under well controlled conditions with sufficient data redundancy, identical peptides across different LC-MS/MS experiments can be compared directly.[22] The quantified proteins/peptides can be mapped onto metabolic pathways and provide insights into the MoA's of these new metallodrugs against diseases.

Herein, the MoA of an osmium-based metallodrug on A2780 ovarian cells and the MoA of an iridium-based metallodrug in *Staphylococcus aureus (S. aureus)* were investigated with nLC-FT-ICR MS/MS.

Experimental

Materials

Chicken lysozyme, formic acid, ammonium bicarbonate (AmmBic), dithiothreitol (DTT), iodoacetamide (IAA), glycerol, acetone, and trypsin were all purchased from Sigma Aldrich (Dorset, England). LCMS grade acetonitrile (MeCN) was purchased from VWR (Lutterworth, UK). Hypersep C18 RP cartridges were purchased from Thermo Fisher Scientific (Waltham, UK). Ultra-pure water (18.2

$M\Omega cm^{-1}$) was obtained from a Milli-Q UV III water purification system (Milli-Q, Hertfordshire, UK).

Synthesis of the osmium-based complex (FY12)

$[(\eta^6\text{-bip})Os(azpy\text{-}NMe_2)I]PF_6$ (FY12) was synthesised and characterised as described previously by Dr. Russell Needham.[23]

Synthesis of the iridium-based complex (CF281)

$[(Cp^{Xbiph})Ir(TolBig)Cl]Cl$ (CF281) was synthesised and characterised as described previously by Dr. Feng Chen.[2]

Preparation of the FY12 - A2780 ovarian cancer cell sample

Approximately 1 x 10^9 of A2780 human ovarian cancer cells were treated with FY12 (IC_{50} concentration, 0.14 μM in 5% DMSO and 95% media). All cells were harvested and lysed by sonicating with glass beads at 277 K in a Tris-buffer solution (50 mM Tris-HCl with 1% glycerol; pH 7.5). The samples were sonicated for 30 s and cooled for 1 min; the sonication process being repeated 45 times for effective cell lysis. All samples were then centrifuged for 30 min at 5,000 g at 277 K to remove the pellets. The supernatant from each sample was then quantified against a standard calibration curve of bovine serum albumin (BSA) using Bradford Reagent (Sigma Aldrich Company Ltd; Dorset, England). The concentrations of proteins in these samples were then calculated using the calibration curve.

The above samples were diluted using 100 mM ammonium bicarbonate solution (Sigma Aldrich Company Ltd; Dorset, England); the disulfide bonds were reduced with 50 mM dithiothreitol (Sigma Aldrich Company Ltd; Dorset, England) and the free cysteine thiols were alkylated with 100 mM iodoacetamide (Sigma Aldrich Company Ltd; Dorset, England). The alkylated lysate was then digested with trypsin (Sigma Aldrich Company Ltd; Dorset, England) by incubating the mixture for 16 h at 310 K. The digested peptides were desalted using C18 SPE cartridges (Thermo Scientific), washed twice with Milli-Q water and then eluted with 80% MeCN/aqueous solution.

Desalted samples were dried down with a Speed Vac concentrator (Thermo Scientific) and re-dissolved in Milli-Q water to a concentration of 1 mg/mL.

Preparation of the CF281 - S. aureus sample

S. aureus was cultured overnight with a culture density of 1 x 10^7, after exposure to the iridium complex at the concentration of 0 and 1 minimum inhibitory concentration (MIC) values (2 samples) for 1 h at 310 K. All cells were harvested and lysed by using a homogeniser at 277 K in a Tris-buffer solution (50 mM Tris-HCl, 2.5% sodium dodecyl sulphate (SDS), and 5% glycerol, pH 7.5). The samples were homogenised for 5 s and cooled down for 30 s at a power setting of 50%; the homogenisation process being repeated 20 times for effective cell lysis. Lysed samples were centrifuged at 13,000 rpm for 20 min at 277 K. Supernatants were transferred into fresh falcon tubes and 5 mL of cold acetone were added to each sample and kept at 253 K for 2 h for protein precipitation. Fractions of samples were collected and allowed to grow on agar plates overnight to ensure all bacterial cells were killed. Samples were centrifuged again at 13,000 rpm for 15 min at 277 K the next morning after ensuring all bacterial cells were killed. Supernatants were removed and fresh cold acetone was added to wash the samples, followed by centrifugation at 13,000 rpm for 15 min at 277 K again. Pellets were transferred into fresh Eppendorfs and 100 µL of 8 M urea was added to re-dissolve the pellets. The re-dissolved pellet was then quantified against a standard calibration curve of bovine serum albumin (BSA) using Bradford reagent (Sigma Aldrich Company Ltd., Dorset, England). The concentrations of proteins in these samples were then back calculated according to the absorbance reading from the microplate reader. The above samples were diluted using 100 mM ammonium bicarbonate solution (Sigma Aldrich Company Ltd., Dorset, England), the disulfide bonds were reduced with 50 mM dithiothreitol (Sigma Aldrich Company Ltd., Dorset, England), and the free cysteine thiols were alkylated with 100 mM iodoacetamide (Sigma Aldrich Company Ltd., Dorset, England). The alkylated lysate was digested with trypsin (1 µg/µL) (Sigma Aldrich Company Ltd., Dorset, England) by incubating the mixture for 16 h at 310 K. The digested peptides were desalted using C18 SPE cartridges (Thermo Scientific), washed twice with Milli-Q water and then eluted with 80:20 MeCN:aqueous solution. Desalted samples were dried down with Speed Vac

concentrator (Thermo Scientific) and re-dissolved in Milli-Q water into concentration of 1 mg/mL.

nLC-FT-ICR MS/MS

nLC separations were achieved using an EASY nano-LC II system (Proxeon, Hemel Hempstead, UK) with a home-made 18 cm, C18 reverse phase (RP) nano capillary column (75 µm, 3 µm particle size, pore size of 300 Å) and a 3 cm C18 RP pre-column (150 µm, 3 µm particle size, pore size of 300 Å). Separation of tryptic cell digest was achieved using a formic acid-acidified water/acetonitrile gradient from 5% MeCN to 30% MeCN over 150 min, holding for 5 minutes at 30% MeCN, followed by a second gradient of 15 min from 30% MeCN to 80% MeCN. Finally, with a 30-min wash of 80% MeCN at a constant flow rate of 500 nL/min.

nLC-MS/MS experiments were carried out by coupling EASY nano-LC II system (Proxeon, Hemel Hempstead, UK) to a Bruker SolariX Fourier Transform-Ion Cyclotron Resonance Mass Spectrometer (FT-ICR MS) fitted with a 12 tesla actively shielded magnet (Bruker Daltonics, Bremen, Germany), *via* a custom nano-spray source utilising glass nanospray tips, 360 µm tip O.D., 50 µm tip I.D., and 8 µm orifice I.D. (New Objective, MA, USA). Singly-charged or unknown-charge species were excluded automatically from the MS/MS list, only the most intense multiply charged ions from each scan were isolated in the quadrupole and accumulated for 1 - 1.2 s in the hexapole collision cell for collisionally activated dissociation (CAD) MS/MS experiments. 1.6 and 2.25 µg (FY12 -A2780 and CF281 - *S. aureus*, respectively) of each sample was injected for each nLC-MS/MS run. All nLC-MS/MS experiments were done with at least 5 technical replicates for consistency, reproducibility, and reliability of the data.

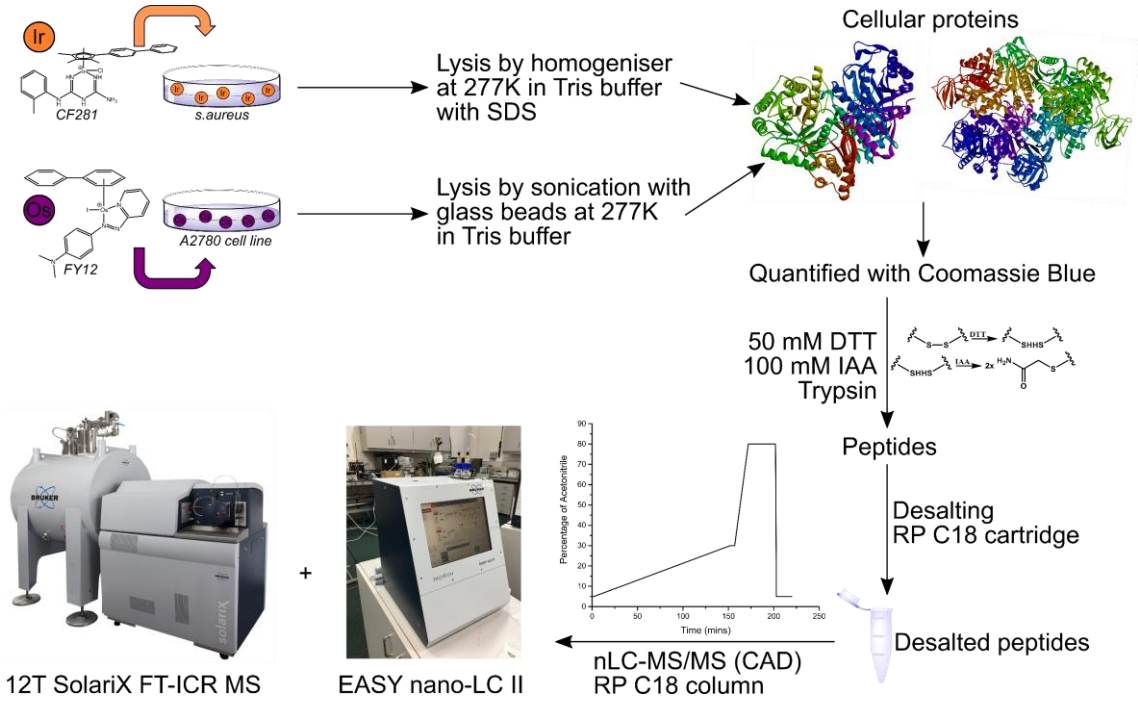


Figure 5.1 nLC-FT-ICR MS/MS workflow of the FY12 - A2780 cancer cell and CF281 - *S. aureus* bacterial cells.

Mass lists were generated by the FTMS peak picking algorithm in DataAnalysis v4.2 (Bruker Daltonik GmbH, Bremen, Germany) from nLC-MS/MS data sets and matched against the Mascot database search engine. All the identified proteins were filtered with 1% false discovery rate (FDR). Each peptide picked out by the search engine was manually analysed from the raw LC-MS/MS data, and any false positive results were critically eliminated. The reduced datasets (after elimination) were then grouped together, and peptides that were detected 3 times or more out of the 5 replicates of data were further selected for quantification.

Quantification of peptides was made by spiking with chicken egg white lysozyme digest (HEWL) to the cancer cell digest in 1:6 HEWL digest : cell digest volume ratio (lysozyme peptides 0.02 μg/μL, cell digest 0.4 μg/μL) into each sample during the nLC-MS/MS runs, and by determining the ratio between the cell peptide of interest and the specific lysozyme peptide. The differences between the ratios from the control sample set and the drug treated sample set were compared. T-tests and p values were determined as measures of the level of significance between the different samples.

Pathway Analysis

Peptides that were detected 3 times or more out of the 5 replicates of data sets were selected for quantification of pathway analysis. The protein entry lists were imported and cross-matched with the Database for Annotation, Visualisation and Integrated Discovery (DAVID) Bioinformation Database,[24] then a list of related pathways was generated and exported by the Kyoto Encyclopedia of Genes and Genomes (KEGG) database.[25–27] Peptides were quantified using the methodology described above. The weighted area ratio was calculated using the equation shown below for proteins with multiple identified peptides. This provided a more comprehensive measure of the change in abundance under different conditions. Protein Clustering was achieved with String database v11.0.

$$Average\ Fold\ of\ Change = \sum(F \times \frac{A}{\sum A}) \qquad \text{[Eq. 1]}$$

where F is the fold of change of a peptide, A is the average area ratio among the 5 replicates of data sets.

Results and Discussion

FY12-A2780 Ovarian Cancer Cells

Two batches of A2780 human ovarian cells (ca 1 x 10^9 each) were cultured, one batch treated with FY12 (10 μM) and the other batch without (control). Proteins from these cells were harvested and digested as described in the experimental section. The digested proteins were analysed by coupling a nLC to the FT-ICR MS equipped with a C18 reversed phase (RP) column. Each of the two samples (control and drug treated) were repeated five times for reproducibility and reliability purposes. These data sets were searched against the human proteome database using the Mascot search engine with the oxidation of amino acid side chains (histidine, tryptophan, and methionine), deamidation of asparagine and glutamine, phosphorylation of serine, threonine, and tyrosine as variable modification and carbamidomethyl (from alkylation) as a fixed modification. With 1% false discovery rate (FDR), all identified peptides were manually checked for false positives. Only peptides which detected three times or more

out of the five replicates under the two different reaction conditions were quantified. 231 unique proteins and ~ 1500 peptides met these requirements and quantified (Figure 5.2).

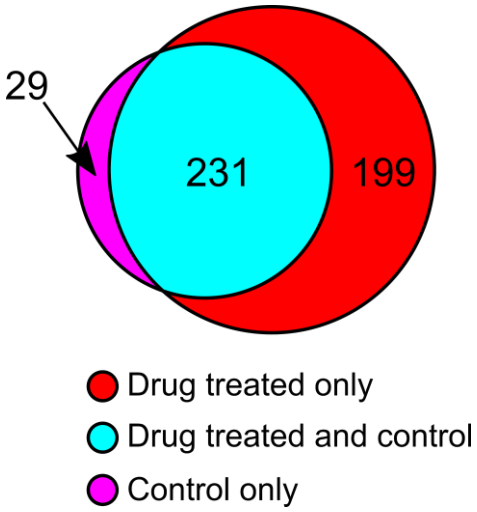


Figure 5.2 Venn diagram of proteins identified by Mascot under two different reaction conditions (control and drug treated).

The 231 protein identities were imported onto the DAVID bioinformatics database (DAVID Bioinformatics Resources 6.8),[24] and mapped onto metabolic pathways, summarised in Figure 5.3.

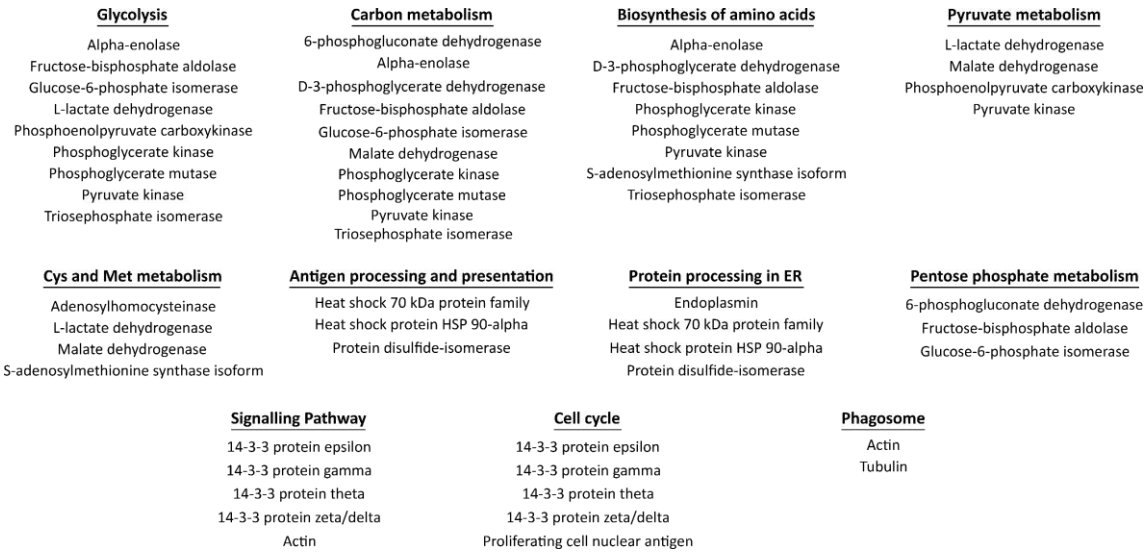


Figure 5.3 Mutual proteins mapped onto eleven different metabolic pathways by DAVID bioinformatics database.

Quantification of peptides was achieved by measuring the peak areas of the peptide of interest and the reference peptide from the extract ion chromatograms (EIC), calculating the ratio and compared between the two conditions.

$$Ion\ Abundance\ Ratio = \frac{A_{peptide\ of\ interest}}{A_{reference\ peptide}} \qquad \text{[Eq. 2]}$$

These mutual proteins were quantified using a label-free approach with lysozyme tryptic digest as an internal control. A specific lysozyme peptide (FESNFNTQATNR, 714.8365 *m/z*) was chosen as the reference peptide and a t-test was performed to ensure the detection of this peptide was statistically consistent enough for quantification purposes between the two different reaction conditions. The p value was calculated to be > 0.05, showing this lysozyme peptide is statistically consistent enough and can be used as a reference peptide for quantification purposes (Table S5.1). The changes in quantity of each protein are summarised in Figure 5.4 and Tables S5.2 and S5.3.

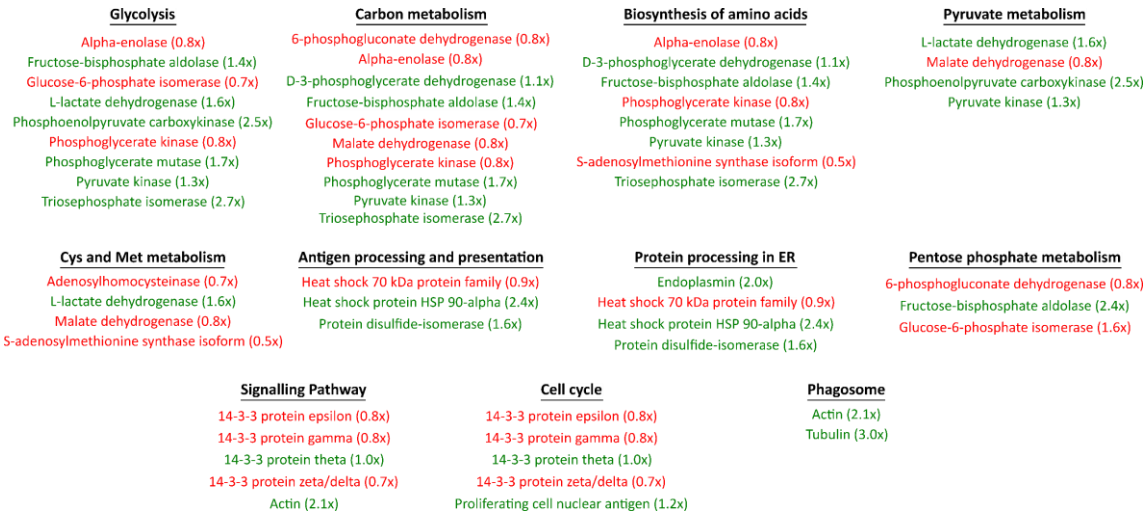


Figure 5.4 Summary of the quantified proteins in metabolic pathways. Quantity changes (fold change) are shown in brackets. Up-regulated proteins are shown in green and down-regulated proteins are shown in red.

Glycolysis pathway was found to have the lowest p value of 1.5×10^{-10}, suggesting this pathway was affected most significantly out of the eleven metabolic pathways (Figure 5.5). Glucose-6-phosphate isomerase (GPI) was found down-

regulated (0.68x) and GPI is responsible for the interconversion of glucose-6-phosphate (α-D-Glucose-6P or β-D-Glucose-6P) and fructose-6-phospate (β-D-Fructose-6P). The down-regulation of GPI prevents the back conversion from β-D-fructose-6P to α/β-D-glucose-6P, supported by the observation of up-regulated fructose-biphosphate aldolase (FBA). FBA was observed to have been up-regulated by 1.37-fold and is responsible for the breakdown of fructose 1,6-biphosphate (β-D-fructose-1,6P2) into dihydroxyacetone (DHAP) and glyceraldehyde 3-phosphate (G3P). Triosephosphate isomerase (TPI 1) was up-regulated by 2.68-fold, the largest change in quantity out of the nine detected proteins in the glycolysis pathway. TPI 1 catalyses the isomerisation of DHAP and G3P, leads to a selective accumulation of G3P in the system.

Phosphoglycerate kinase was down-regulated (PGK, 0.83x) while bisphosphoglycerate mutase was up-regulated (BPGM, 1.66x). PGK contributes to the interconversion of 2-phosphoglycerate (glycerate-2P) and 3-phosphoglycerate (glycerate-3P). While BPGM is responsible for the conversion of 2,3-bisphosphoglycerate (glycerate-2,3P2) from 1,3-bisphosphoglycerate (glycerate-1,3P2) and the formation of glycerate-2P, which is an irreversible reaction. glycerate-2P serves as a substrate in the ninth step of the glycolysis pathway and plays an important role in the eventual production of pyruvate.

α-Enolase (AE) contributes to the interconversion between glycerate-2P and phosphoenol pyruvate (PEP), which is a reversible reaction. When A2780 ovarian cancer cells were treated with FY12, AE was down-regulated by 0.77-fold, while phosphoenolpyruvate carboxykinase (PCK 1,2) was up-regulated by 2.46-fold, the second largest change upon drug treatment. PCK 1,2 irreversibly converts oxaloacetate into PEP. With the down-regulation of AE and up-regulation of PCK 1,2, the amount of PEP was vastly accumulated selectively. Pyruvate kinase (PK) and lactate dehydrogenase (L-LDH) were both up-regulated, by 1.34-fold and 1.58-fold, respectively. These two enzymes contribute to the production of pyruvate, which supplies energy to cells through the citrate cycle, also known as the Krebs cycle when oxygen is present.[28] The selective accumulation of pyruvate is believed to result from a combination of uncontrolled proliferation and stress response of the cancer cells. The

uncontrolled proliferation requires a large energy supply for cell replication, but more importantly, the quantity change was achieved by comparing the control and the drug treated samples, and these changes were induced by FY12. When cancer cells are present under a stressful condition/environment or dying, they would adapt and undergo more replication/reproduction, as a secondary response.

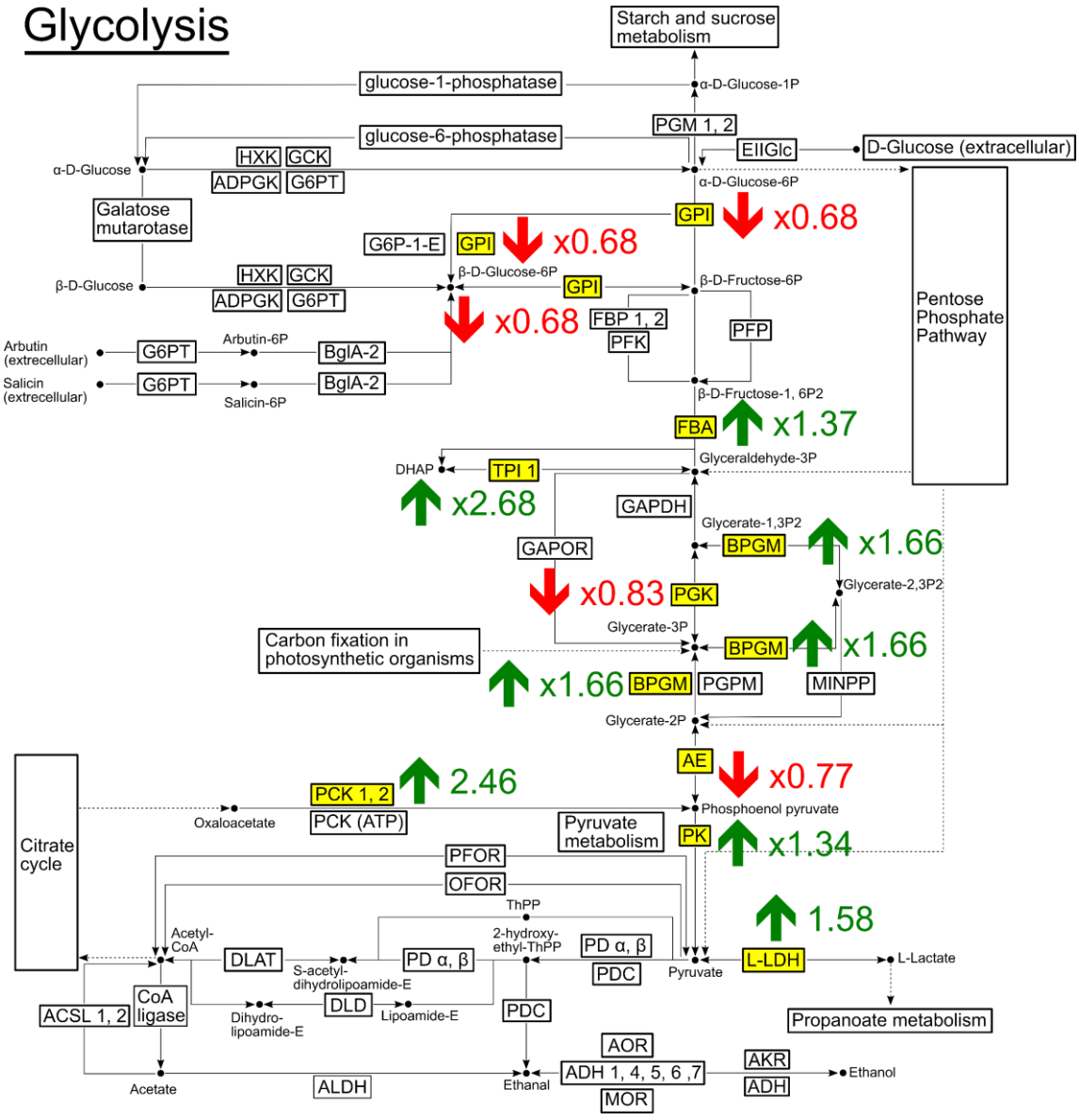


Figure 5.5 Effects of FY12 on the level of proteins associated with the glycolysis pathway in A2780 ovarian cancer cells. Nine proteins highlighted in yellow were identified by nLC-CAD MS/MS, with the magnitude of the fold-change indicated in green (up-regulated) or red (down-regulated).

Key: **PGM**: Phosphoglucomutase; **EIIGlc**: D-glucose Npi-phosphotransferase; **HXK**: Hexokinase; **GCK**: Glucokinase; **ADPGK**: ADP dependent glucokinase; **G6PT**: D-glucose-6-phosphotransferase; **G6P-1-E**: D-glucose-6-phosphate 1-epimerase; **GPI**: Glucose-6-phosphate isomerase; **BgIA-2**: 6-phospho-beta-D-glucosyl-D-glucose glycohydrolase; **FBP**: Fructose-bisphosphatase; **PFK**: Phosphofructokinase; **PFP**: D-fructose-6-Phosphate 1-phosphotransferase; **FBA**: Fructose-bisphosphate aldolase; **TPI**: Triosephosphate isomerase; **GAPDH**: Glyceraldehyde-3-phosphate dehydrogenase; **GAPOR**: D-glyceraldehyde-3-phosphate: NADP+ oxidoreductase; **BPGM**: Bisphosphoglycerate mutase; **PGK**: Phosphoglycerate kinase; **PGPM**: D-phosphoglycerate 2,3-phosphomutase; **MINPP**: Multiple inositol-polyphosphate phosphatase; **AE**: Alpha-enolase; **PCK**: Phosphoenolpyruvate carboxykinase; **PCK (ATP)**: ATP Oxaloacetate carboxy-lyase; **PK**: pyruvate kinase; **PFOR**: Pyruvate ferredoxin-2-oxidoreductase; **OFOR**: 2-oxocarboxylate: ferredoxin-2-oxidoreductase; **ACSL**: Acyl-CoA-synthetase; **DLAT**: Dihydrolipoamide S-acetyltransferase; **DLD**: Dihydrolipoamide dehydrogenase; **PD**: Pyruvate dehydrogenase; **PDC**: 2-oxo-acid-carboxyl-lyase; **L-LDH**: Lactate dehydrogenase; **ALDH**: Aldehyde dehydrogenase; **ADH**: Alcohol dehydrogenase; **AOR**: Alcohol cytochrome c oxidoreductase; **MOR**: Methanol cytochrome c oxidoreductase; **AKR**: Aldo-keto reductase.

As the same protein can be involved in more than one pathway, the quantified proteins were also imported onto the String database for clustering purposes. This allows a clearer observation of how these proteins are related to each other and how are they contributing in the cells.

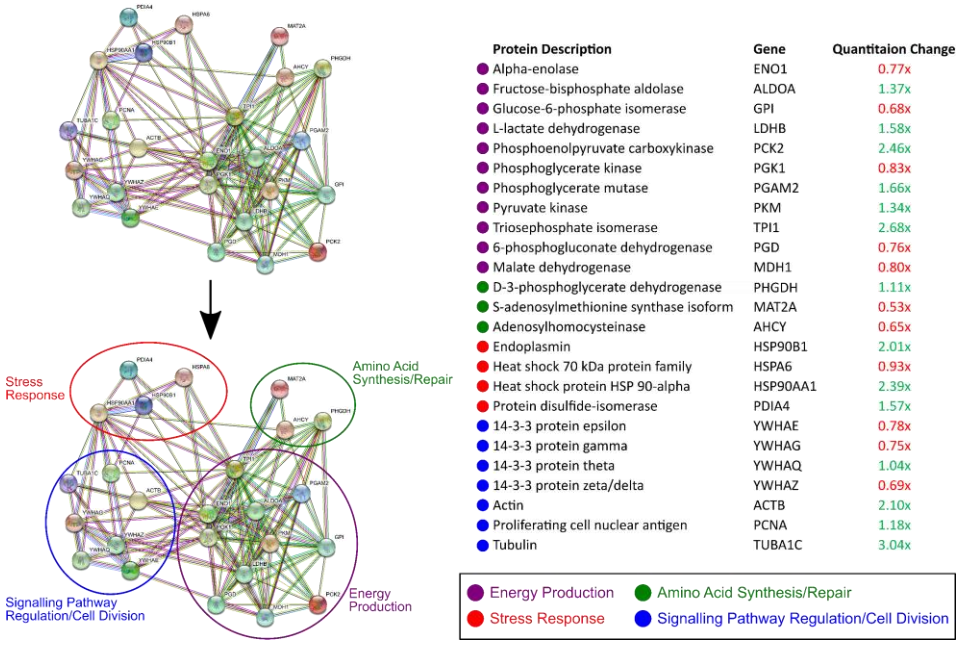
25 proteins were imported onto String database, and they were grouped into 4 major clusters according to their cellular functions, including: (1) signalling pathway/cell division, (2) stress response, (3) amino acid synthesis/repair, and (4) energy production (Figure 5.6).

When A2780 ovarian cancer cells were treated with FY12, the cancer cells were under a stressful environment, which stimulated a lot of primary and secondary responses. Three of the four 14-3-3 protein isoforms were down-regulated and they have been reported to be closely related to cancers and play important roles in many

signalling pathways.[29] 14-3-3 protein epsilon (14-3-3 ε) is one of the mammalian 14-3-3 protein family members that is highly conserved in eukaryotes and has been reported as an oncogene activator, with the potential to cause cancer. In tumour cells, oncogenes are often overexpressed and cause the tumour cells designated for apoptosis to survive and proliferate instead.[30] A down-regulation (0.78-fold) of 14-3-3 ε is believed to leave the tumour cells scheduled to apoptosis/programmed cell death. 14-3-3 protein gamma (14-3-3 γ) was found to be involved in cancerous processes, and is consistently up-regulated in other cancers such as human lung cancer.[31] The observation of down-regulation (0.75-fold) of this protein may suggest FY12 has a positive impact in killing the ovarian cancer cells. Overexpression of 14-3-3 protein zeta (14-3-3 ζ) was previously reported as contributing apoptosis resistance, cancer recurrence and chemoresistance.[32] 14-3-3 ζ can therefore be used as a target or biomarker for cancer therapeutic resistance. The down-regulation of 14-3-3 ζ by 0.69-fold shows FY12 is effectively targeting an important metabolic pathway and blocks the function of 14-3-3 ζ in human cancers, lowering the opportunity of having cancer recurrence and chemoresistance, increasing the survival rate of patients. Actin and tubulin were found up-regulated by 2.1-fold and 3.04-fold respectively, the two largest quantity changes observed. These two proteins are heavily involved in the cell division stages/cell proliferation. Actin contributes to cell motility/invasivity of cells and cell division. When the ovarian cancer cells were "attacked" by the FY12 osmium-based anticancer drug, the cancer cells were under a vast amount of stress and appear to have stimulated a secondary response of triggering more cell division/replication. Tubulin is heavily involved in mitosis, especially for migrating chromosomes to the opposite poles of a mitosing cells during the anaphase. Similar to actin, the up-regulation of tubulin is also believed to be a secondary response of from the stimulated cancer cell division.

Heat shock proteins are a family of proteins that are produced by cells in response to exposure to stressful conditions, such as heat,[33] and UV light.[34] They are chaperone proteins that assist other proteins to fold properly and possess specific functionality. Endoplasmin (HSP90B1) and heat shock protein HSP 90-alpha (HSP90AA1) were found up-regulated by 2.01-fold and 2.39-fold, respectively, which indicates that FY12 induced stress onto the ovarian cancer cells, and the heat shock

proteins were responding by up-regulating these proteins to cope with the stressful environment, possibly by altering the folding of proteins. This is supported by the up-regulation of the protein disulfide-isomerase, which is an enzyme that catalyses the formation and cleavage of cystine disulfide bonds within proteins as they fold.[35] Healthy biological systems have stable amino acid synthesis, in order to achieve a sustainable protein production with correct folding and functionalities. However, once the delicate balance of amino acid synthesis is broken, proteins can be folded differently causing destructive effects. S-Adenosylmethionine synthase isoform (MAT2A) and adenosylhomocysteinase (AHCY) were down-regulated by 0.53-fold and 0.65-fold, respectively. The synthesis of S-adenosylmethionine is often referred to as the rate-determining step of the methionine synthesis cycle.[36] The decrease in quantity limits the production of methionine and possibly disrupts important cellular functions such as the growth of new blood vessels (angiogenesis). ACHY is an enzyme that converts adensylhomocysteine to homocysteine and adenosine. Similar to the decreased quantity of MAT2A, the decreased quantity of ACHY leads to less cysteine production, which affects protein folding greatly, as the cystine disulfide bonds contribute significantly to protein folding.[37]



Protein Description	Gene	Quantitaion Change
Alpha-enolase	ENO1	0.77x
Fructose-bisphosphate aldolase	ALDOA	1.37x
Glucose-6-phosphate isomerase	GPI	0.68x
L-lactate dehydrogenase	LDHB	1.58x
Phosphoenolpyruvate carboxykinase	PCK2	2.46x
Phosphoglycerate kinase	PGK1	0.83x
Phosphoglycerate mutase	PGAM2	1.66x
Pyruvate kinase	PKM	1.34x
Triosephosphate isomerase	TPI1	2.68x
6-phosphogluconate dehydrogenase	PGD	0.76x
Malate dehydrogenase	MDH1	0.80x
D-3-phosphoglycerate dehydrogenase	PHGDH	1.11x
S-adenosylmethionine synthase isoform	MAT2A	0.53x
Adenosylhomocysteinase	AHCY	0.65x
Endoplasmin	HSP90B1	2.01x
Heat shock 70 kDa protein family	HSPA6	0.93x
Heat shock protein HSP 90-alpha	HSP90AA1	2.39x
Protein disulfide-isomerase	PDIA4	1.57x
14-3-3 protein epsilon	YWHAE	0.78x
14-3-3 protein gamma	YWHAG	0.75x
14-3-3 protein theta	YWHAQ	1.04x
14-3-3 protein zeta/delta	YWHAZ	0.69x
Actin	ACTB	2.10x
Proliferating cell nuclear antigen	PCNA	1.18x
Tubulin	TUBA1C	3.04x

Figure 5.6 FY12-A2780 ovarian cancer cell experiment. Protein clustering according to their cellular functions, achieved with String v11.0.

CF281-Staphylococcus aureus Bacterial Cells

The same proteomics approach was performed to study the effect of the iridium-based antimicrobial complex on *S. aureus* bacterial cells. Each of the two samples (control and dosed with 1 MIC of iridium) were repeated five times for reproducibility and reliability purposes. These data sets were searched against the *S. aureus* proteome database using the Mascot search engine with the oxidation of amino acid side chains (histidine, tryptophan, and methionine), deamidation of asparagine and glutamine, phosphorylation of serine, threonine, and tyrosine as variable modification and carbamidomethyl (from cysteine alkylation) as a fixed modification. With 1% false discovery rate (FDR), all identified peptides were manually checked for false positives. Only peptides which were detected three times or more out of the five replicates under the two different reaction conditions were quantified. 81 unique proteins and ~ 300 peptides met these requirements and quantified (Figure 5.7).

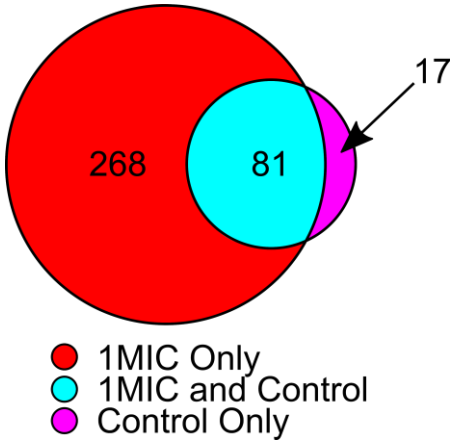


Figure 5.7 Venn diagram of proteins identified by Mascot under two different reaction conditions (control and 1 MIC).

The difference in the number of proteins detected in the two reaction conditions was significantly different (~ 3.6 times more compared to the control sample). The overall protein content after the cell lysis was quantified using Bradford reagent and measurement of the light absorbance at 600 nm. Proteins turn the Bradford reagent from brown to blue, the more concentrated the protein is, the darker blue. However, sodium dodecyl sulphate (SDS) which was used as a detergent to break down the

bacterial cell wall, will turns the Bradford reagent from brown to a greenish blue colour. In theory, all SDS should have been removed using acetone precipitation, however the residual SDS could induce artefacts into the protein quantification using light absorbance measurements by absorbing the 600 nm wavelength irradiation too. In order to obtain a measurable absorbance difference, the colour change of Bradford reagent must be significantly high enough. However, compared to the drug-treated sample, which required only 1 µL of the sample to cause a significant change, the control sample required 5 µL. Therefore, the apparent protein quantity of the control sample may have appeared to be larger while a proportion of the light absorbance was contributed by the SDS, causing a low efficiency extraction for the Bradford assay calibration using bovine serum albumin (BSA). The experiment could be improved by increasing the initial number of bacterial cells, in order to overcome the concentration of residual SDS from the cell lysis procedures. The increase in number of cells could also avoid any scaling during sample preparation or quantification stages, yielding a more reliable protein quantity.

All the peptides were quantified using the label-free quantification method with the lysozyme peptide (NTDGSTDYGILQINSR, 877.4212 *m/z*) as an internal control. A t-test was performed on the internal control, to ensure the detected abundances are statistically similar between the control and the drug treated sample (1 MIC). The p value was calculated to be 0.12 (> 0.05), which shows the detected abundances were similar and can be used as an internal control for quantification purposes (Table S5.4). Peptides were quantified using the same method as mentioned above and only peptides with p values from t-test lower than 0.05 when compared between the two reaction conditions were selected for protein clustering (Table S5.5). 26 unique proteins met the requirements and were imported into the String database v11.0 for protein clustering. The proteins were separated into 6 groups according to their cellular functions (Figure 8), including: (1) stress response, (2) energy production, (3) bacteria self-defence, (4) stress/repair, (5) replication, and (6) stress and removal of drugs.

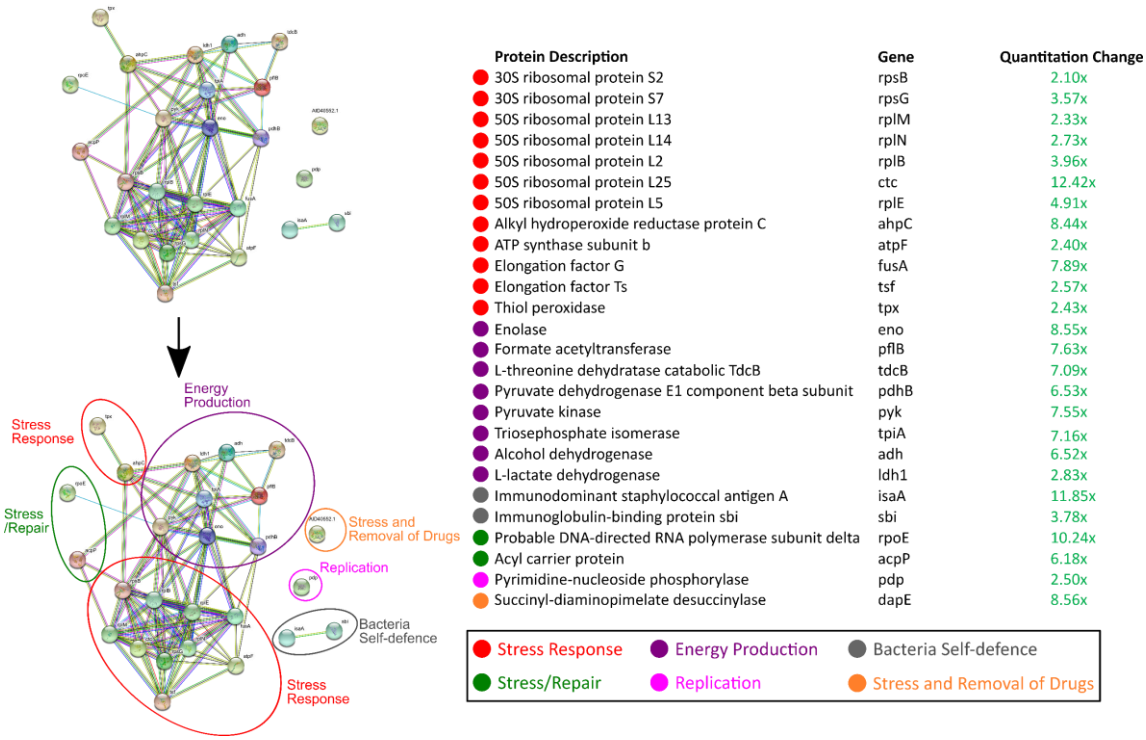


Figure 5.8 CF281-*S. aureus* bacterial cell experiment. Protein clustering according to their cellular functions, achieved with String v11.0.

Analysis of the data suggest that when *S. aureus* bacterial cells were treated with CF281, a lot of stress was induced as a secondary response. Different subunits of 30S ribosomal protein and 50S ribosomal protein were found up-regulated, ranging from 2.1-fold to 12.4-fold, with the quantity of the 50S ribosomal protein L25 subunit increased the most. These ribosomal subunits are heavily involved in the translation process, synthesising proteins after the transcription of DNA to RNA.[38] When the bacterial cells are under a stressful environment, they are likely to be stimulated to produce more proteins for different purposes, such as enzymes for energy production, or for cell replication. Apart from the ribosomal proteins, two elongation factors, G and Ts, were up-regulated by 7.9-fold and 2.6-fold, respectively. They are mainly involved in protein translation, by catalysing the movement of translation RNA (tRNA) and messenger RNA (mRNA) through ribosome. Alkyl hydroperoxide reductase C (ahpC) and thiol peroxidase (tpx) were also up-regulated by 8.4-fold and 2.4-fold, respectively. Bacteria, like cancer cells, are constantly under high oxidative stress,[39] therefore, as an evolution, bacterial cells have developed a strong antioxidant response to reactive

oxygen species (ROS).[40] Proteins ahpC and tpx both play important roles in cell protection against oxidative stress. The observation of up-regulation between the control and drug-treated samples suggest that the iridium-based antimicrobial complex induced even more oxidative stress into the bacteria and triggered the oxidative stress defence system. Therefore, inducing a vast amount of oxidative stress could be important to the MoA of CF281 (Figure 5.8, highlighted in red).

Eight proteins were found involved in the energy production process, and they were all up-regulated, ranging from 2.8-fold to 8.6-fold, with the highest increase for enolase (Figure 8, highlighted in purple). Enolase is responsible for the conversion of glycerate-2P to phosphoenol pyruvate (PEP), the ninth and penultimate step of glycolysis. Formate acetyltransferase (pflB) helps regulate anaerobic glucose metabolism,[41] and was also up-regulated, by 7.6-fold. *S. aureus* are facultative anaerobes, which means they generate adenosine triphosphate (ATP) by aerobic respiration if oxygen is present, but are also capable of switching to fermentation or anaerobic respiration if they are in a hypoxia environment. These observations support the secondary responses after the drug treatment, such as extra protein production and cell replication.

Immunodominant staphylococcal antigen A (IsaA) is a non-covalently cell wall-bound lytic transglycosylase, and involved in cell wall growth and division.[42] Immunoglobulin-binding protein sbi (sbi) has been reported to protect the bacterial cell against the host immune response.[43] IsaA and sbi were both up-regulated by 11.9-fold and 3.8-fold, respectively. This is believed to be a secondary response from the bacterial cells, responding to the "threat" introduced by CF281, as a self-defence system (Figure 5.8, highlighted in grey).

Acyl carrier protein (AcpP) is an important component in both fatty acid and polyketide biosynthesis, many polyketides have been reported to possess antimicrobial properties, such as erythromycin A.[44] RNA polymerase (rpoE) is involved heavily in the DNA transcription process. RNA polymerase unzips the DNA double strand into single strand and copy them into RNA sequences. They were up-regulated by 10.2-fold and 6.2-fold (Figure 5.8, highlighted in green), as a secondary response from the drug

treatment. Pyrimidine-nucleoside phosphorylase was also found up-regulated by 2.5-fold and it is responsible to synthesise pyrimidine bases for DNA replication. Succinyl-diaminopimelate desuccinylase (dapE) is closely related to the multidrug export system.[45] Multiple drug resistance (MDR) is antimicrobial resistance shown by a species of microorganism to multiple antimicrobial drugs. dapE also catalyses the hydrolysis of N-succinyl-L,L-diaminopimelic acid (SDAP), forming succinate and L,L-2,6-diaminoheptanedioate (DAP), which is an intermediate involved in the biosynthesis of meso-diaminopimelic acid, an essential component of bacterial cell wall in gram-positive bacteria. The up-regulation of dapE (8.6-fold) is believed to also be a secondary response from the antimicrobial treatment.

Conclusions

Pathway analysis of an osmium-based metallodrug on A2780 ovarian cancer cells was achieved by nLC-FT-ICR CAD MS/MS. The protein identities were achieved by matching the fragmentation patterns against the Mascot database and quantification of these proteins achieved using a label-free approach with lysozyme tryptic digested peptides. FY12 has shown promising effects on multiple metabolic pathways and different cellular functions, such as energy production from the glycolysis pathway. Nine unique proteins showed a significant change in quantity (up-regulated) and pyruvate was selectively accumulated for extra energy production. Other than the production of energy, the quantity of 14-3-3 protein isoforms were also influenced when the cancer cells were treated with FY12, which is believed to be the primary MoA against cancer cells. Different stress responses, including up-regulation of heat shock proteins and the down-regulation of amino acid synthesis were believed to be a secondary response from the cancer cells. Previous studies on the MoA's of FY12 and other osmium-based complexes suggested oxidative stress was induced near the mitochondria of the target cells by high superoxide concentration, causing mitochondrial dysfunction and DNA damage.[3,4] The results reported here not only supported these previous studies but also provided a more in-depth investigation in the quantity change of specific proteins.

The mechanism of action (MoA) of the iridium-based antimicrobial complex CF281 was investigated using a similar methodology. 81 unique proteins were detected three times or more out of the replicates, and 26 of those were shown to have significant up-regulated quantity change with p values less than 0.05 in t-tests. These proteins were clustered together according to their cellular functions. The bacterial cells were shown to be in a very stressful environment when they were treated with CF281 and triggered a series of secondary responses, including extra protein/DNA biosynthesis and energy production. Proteins responsible for coping with oxidative stress were also found vastly up-regulated, suggesting CF281 is capable to induce extra oxidative stress inside the bacteria cells and trigger cell death, different MoA from traditional treatments, such as penicillin, which focus on the disruption of cell wall formation by inhibiting the formation of peptidoglycan cross linkages that provide rigidity and strength of the cell wall. The difference in MoA to traditional antimicrobial treatments will provide a possibility to improve the multiple drug resistance problem.

References

(1) Llor, C.; Bjerrum, L. Antimicrobial Resistance : Risk Associated with Antibiotic Overuse and Initiatives to Reduce the Problem. *Ther. Adv. Drug Sade.* **2014**, *5* (6), 229–241.

(2) Chen, F.; Moat, J.; McFeely, D.; Clarkson, G.; Hands-Portman, I. J.; Furner-Pardoe, J. P.; Harrison, F.; Dowson, C. G.; Sadler, P. J. Biguanide Iridium(III) Complexes with Potent Antimicrobial Activity. *J. Med. Chem.* **2018**, *61* (16), 7330–7344.

(3) Hearn, J. M.; Romero-Canelón, I.; Munro, A. F.; Fu, Y.; Pizarro, A. M.; Garnett, M. J.; McDermott, U.; Carragher, N. O.; Sadler, P. J. Potent Organo-Osmium Compound Shifts Metabolism in Epithelial Ovarian Cancer Cells. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112* (29), E3800–E3805.

(4) Hearn, J. M.; Hughes, G. M.; Romero-Canelón, I.; Munro, A. F.; Rubio-Ruiz, B.; Liu, Z.; Carragher, N. O.; Sadler, P. J. Pharmaco-Genomic Investigations of Organo-Iridium Anticancer Complexes Reveal Novel Mechanism of Action. *Metallomics* **2018**, *10* (1), 93–107.

(5) Nam, J. S.; Kang, M. G.; Kang, J.; Park, S. Y.; Lee, S. J. C.; Kim, H. T.; Seo, J. K.; Kwon, O. H.; Lim, M. H.; Rhee, H. W.; et al. Endoplasmic Reticulum-Localized Iridium(III) Complexes as Efficient Photodynamic Therapy Agents *via* Protein Modifications. *J. Am. Chem. Soc.* **2016**, *138* (34), 10968–10977.

(6) Zhang, P.; Chiu, C. K. C.; Huang, H.; Lam, Y. P. Y.; Habtemariam, A.; Malcomson, T.; Paterson, M. J.; Clarkson, G. J.; O'Connor, P. B.; Chao, H.; et al. Organoiridium Photosensitizers Induce Specific Oxidative Attack on Proteins within Cancer Cells. *Angew. Chemie Int. Ed.* **2017**, 14898–14902.

(7) Wootton, C. A.; Sanchez-Cano, C.; Liu, H.-K.; Barrow, M. P.; Sadler, P. J.; O'Connor, P. B. Binding of an Organo-Osmium(II) Anticancer Complex to Guanine and Cytosine on DNA Revealed by Electron-Based Dissociations in High Resolution Top-Down FT-ICR Mass Spectrometry. *Dalton Trans.* **2015**, *44*

(8), 3624–3632.

(8) Wu, K.; Hu, W.; Luo, Q.; Li, X.; Xiong, S.; Sadler, P. J.; Wang, F. Competitive Binding Sites of a Ruthenium Arene Anticancer Complex on Oligonucleotides Studied by Mass Spectrometry: Ladder-Sequencing versus Top-Down. *J. Am. Soc. Mass Spectrom.* **2013**, *24* (3), 410–420.

(9) Wills, R. H.; Habtemariam, A.; Lopez-Clavijo, A. F.; Barrow, M. P.; Sadler, P. J.; O'Connor, P. B. Insights into the Binding Sites of Organometallic Ruthenium Anticancer Compounds on Peptides Using Ultra-High Resolution Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2014**, *25* (4), 662–672.

(10) Gabbiani, C.; Casini, A.; Kelter, G.; Cocco, F.; Cinellu, M. A.; Fiebig, H.-H.; Messori, L. Mechanistic Studies on Two Dinuclear Organogold(III) Compounds Showing Appreciable Antiproliferative Properties and a High Redox Stability. *Metallomics* **2011**, *3* (12), 1318–1323.

(11) Martins, A. P.; Marrone, A.; Ciancetta, A.; Cobo, A. G.; Echevarría, M.; Moura, T. F.; Re, N.; Casini, A.; Soveral, G. Targeting Aquaporin Function: Potent Inhibition of Aquaglyceroporin-3 by a Gold-Based Compound. *PLoS One* **2012**, *7* (5).

(12) Gras, M.; Therrien, B.; Süss-Fink, G.; Casini, A.; Edafe, F.; Dyson, P. J. Anticancer Activity of New Organo-Ruthenium, Rhodium and Iridium Complexes Containing the 2-(Pyridine-2-Yl)Thiazole N,N-Chelating Ligand. *J. Organomet. Chem.* **2010**, *695* (8), 1119–1125.

(13) Banerjee, S.; Soldevila-Barreda, J. J.; Wolny, J. A.; Wootton, C. A.; Habtemariam, A.; Romero-Canelon, I. New Activation Mechanism for Half-Sandwich Organometallic Anticancer Complexes. *Chem. Sci.* **2018**, *9*, 3177–3185.

(14) Wootton, C. A.; Sanchez-Cano, C.; Lopez-Clavijo, A. F.; Shaili, E.; Barrow, M. P.; Sadler, P. J.; O'Connor, P. B. Sequence-Dependent Attack on Peptides by Photoactivated Platinum Anticancer Complexes. *Chem. Sci.* **2018**, *9* (10), 2733–

2739.

(15) Shi, H.; Wang, Q.; Venkatesh, V.; Feng, G.; Young, L. S.; Romero-Canelon, I.; Zeng, M.; Sadler, P. J. Photoactive Platinum(IV) Complex Conjugated to a Cancer-Cell-Targeting Cyclic Peptide. *Dalton Trans.* **2019**, *48*, 8560–8564.

(16) Josefsen, L. B.; Boyle, R. W. Photodynamic Therapy and the Development of Metal-Based Photosensitisers. *Met. Based. Drugs* **2008**, *2008*, 1–23.

(17) Farrer, N. J.; Salassa, L.; Sadler, P. J. Photoactivated Chemotherapy (PACT): The Potential of Excited-State d-Block Metals in Medicine. *Dalton Trans.* **2009**, No. 48, 10690–10701.

(18) Jennings, K. R. Colllision-Induced Decompositions of Aromatic Ions. *Int. J. Mass Spectrom. Ion Phys.* **1968**, *1* (3), 227–235.

(19) Zubarev, R.; Kelleher, N. L.; McLafferty, F. W. Electron Capture Dissociation of Multiply Charged Protein Cations. A Nonergodic Process. *J. Am. Chem. Soc* **1998**, *120* (16), 3265–3266.

(20) Little, D. P.; Speir, J. P.; Senko, M. W.; O'Connor, P. B.; McLafferty, F. W. Infrared Multiphoton Dissociation of Large Multiply Charged Ions for Biomolecule Sequencing. *Anal. Chem.* **1994**, *66* (18), 2809–2815.

(21) Theisen, A.; Yan, B.; Brown, J. M.; Morris, M.; Bellina, B.; Barran, P. E. Use of Ultraviolet Photodissociation Coupled with Ion Mobility Mass Spectrometry to Determine Structure and Sequence from Drift Time Selected Peptides and Proteins. *Anal. Chem.* **2016**, *88* (20), 9964–9971.

(22) Wong, J. W. H.; G., C. *An Overview of Label-Free Quantification Methods in Proteomics by Mass Spectrometry*; 2009.

(23) Fu, Y.; Habtemariam, A.; Pizarro, A. M.; Van Rijt, S. H.; Healey, D. J.; Cooper, P. A.; Shnyder, S. D.; Clarkson, G. J.; Sadler, P. J. Organometallic Osmium Arene Complexes with Potent Cancer Cell Cytotoxicity. *J. Med. Chem.* **2010**, *53* (22), 8192–8196.

(24) Huang, D. W.; Sherman, B. T.; Lempicki, R. A. Systematic and Integrative Analysis of Large Gene Lists Using DAVID Bioinformatics Resources. *Nat. Protoc.* **2009**, *4* (1), 44–57.

(25) Kanehisa, M.; Goto, S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* **2000**, *2844* (1), 27–30.

(26) Kanehisa, M.; Sato, Y.; Kawashima, M.; Furumichi, M.; Tanabe, M. KEGG as a Reference Resource for Gene and Protein Annotation. *Nucleic Acids Res.* **2016**, *44* (D1), D457–D462.

(27) Kanehisa, M.; Furumichi, M.; Tanabe, M.; Sato, Y.; Morishima, K. KEGG: New Perspectives on Genomes, Pathways, Diseases and Drugs. *Nucleic Acids Res.* **2017**, *45* (D1), D353–D361.

(28) Gray, L. R.; Tompkins, S. C.; Taylor, E. B. Regulation of Pyruvate Metabolism and Human Disease. *Cell. Mol. Life Sci.* **2014**, *71* (14), 2577–2604.

(29) Thomas, D.; Guthridge, M.; Woodcock, J.; Lopez, A. 9 - 14-3-3 Protein Signaling in Development and Growth Factor Responses. In *Current Topics in Developmental Biology*; Schatten, G. P., Ed.; Current Topics in Developmental Biology; Academic Press, 2005; Vol. 67, pp 285–303.

(30) Leal, M. F.; Calcagno, D. Q.; Demachki, S.; Assumpção, P. P.; Chammas, R. Clinical Implication of 14-3-3 Epsilon Expression in Gastric Cancer. *World J. Gastroenterol.* **2012**, *18* (13), 1531–1537.

(31) Qi, W.; Liu, X.; Qiao, D.; Martinez, J. D. Isoform-Specific Expression of 14-3-3 Proteins in Human Lung Cancer Tissues. *Int. J. Cancer* **2005**, *113* (3), 359–363.

(32) CL, N.; D, Y. 14-3-3ζ as a Prognostic Marker and Therapeutic Target for Cancer. *Expert Opin Their Targets* **2010**, *14* (12), 1343–1354.

(33) Ritossa, F. A New Puffing Pattern Induced by Temperature Shock and DNP in Drosophila. *Cell. Mol. Life Sci.* **1962**, *18* (12), 571–573.

(34) Cao, Y.; Ohwatari, N.; Matsumoto, T.; Kosaka, M.; Ohtsuru, A.; Yamashita, S.

TGF-B1 Mediates 70-KDa Heat Shock Protein Induction Due to Ultraviolet Irradiation in Human Skin Fibroblasts. *Pflugers Arch. Eur. J. Physiol.* **1999**, *438* (3), 239–244.

(35) Wilkinson, B.; Gilbert, H. F. Protein Disulfide Isomerase. *Biochim. Biophys. Acta - Proteins Proteomics* **2004**, *1699* (1–2), 35–44.

(36) Markham, G. D.; Pajares, M. A. Structure-Function Relationships in Methionine Adenosyltransferases. *Cell. Mol. Life Sci.* **2009**, *66* (4), 636–648.

(37) Netto, L. E. S.; de Oliveira, M. A.; Monteiro, G.; Demasi, A. P. D.; Cussiol, J. R. R.; Discola, K. F.; Demasi, M.; Silva, G. M.; Alves, S. V.; Faria, V. G.; et al. Reactive Cysteine in Proteins: Protein Folding, Antioxidant Defense, Redox Signaling and More. *Comp. Biochem. Physiol. - C Toxicol. Pharmacol.* **2007**, *146* (1-2 SPEC. ISS.), 180–193.

(38) Xiang, Z.; Wen-Juan, L.; Jun-Ming, L.; Peng, L.; Hua, L. Ribosomal Proteins: Functions beyond the Ribosome. *J. Mol. Cell Biol.* **2015**, *7* (2), 92–104.

(39) Dharmaraja, A. T. Role of Reactive Oxygen Species (ROS) in Therapeutics and Drug Resistance in Cancer and Bacteria. *J. Med. Chem.* **2017**, *60*, 3221–3240.

(40) Imlay, J. A. Cellular Defenses against Superoxide and Hydrogen Peroxide. *Annu. Rev. Biochem.* **2008**, *77* (1), 755–776.

(41) Böck, A.; Sawers, G. Anaerobic Regulation of Pyruvate Formate-Lyase from Escherichia Coli K-12. *J. Bacteriol.* **1988**, *170* (11), 5330–5336.

(42) Lorenz, U.; Lorenz, B.; Schmitter, T.; Streker, K.; Erck, C.; Wehland, J.; Nicke, J.; Zimmermann, B.; Ohlsen, K. Functional Antibodies Targeting IsaA of Staphylococcus Aureus Augment Host Immune Response and Open New Perspectives for Antibacterial Therapy. *Antimicrob. Agents Chemother.* **2011**, *55* (1), 165–173.

(43) Zhang, L.; Jacobson, K.; Vasi, J.; Lindberg, M.; Frykberg, L. A Second IgG-Binding Protein in Staphylococcus Aureus. *Microbiology* **1998**, *144*, 985–991.

(44) Khan, N. W.; Hassan, F.; Naqvi, B. S.; Hasan, S. M. F. Antimicrobial Activity of Erythromycin and Clarithromycin against Clinical Isolates of Escherichia Coli, Staphylococcus Aureus, Klebsiella and Proteus by Disc Diffusion Method. *Pak. J. Pharm. Sci.* **2011**, *24* (1), 25–29.

(45) Projan, S. J.; Luong, T. T.; Dunman, P. M.; Lee, C. Y.; Murphy, E. Transcription Profiling of the MgrA Regulon in Staphylococcus Aureus. *J. Bacteriol.* **2006**, *188* (5), 1899–1910.

Supplementary Information

Figure S5.1 Example of base peak chromatogram of A2780 ovarian cancer cell control nLC-FT-ICR CAD MS/MS.

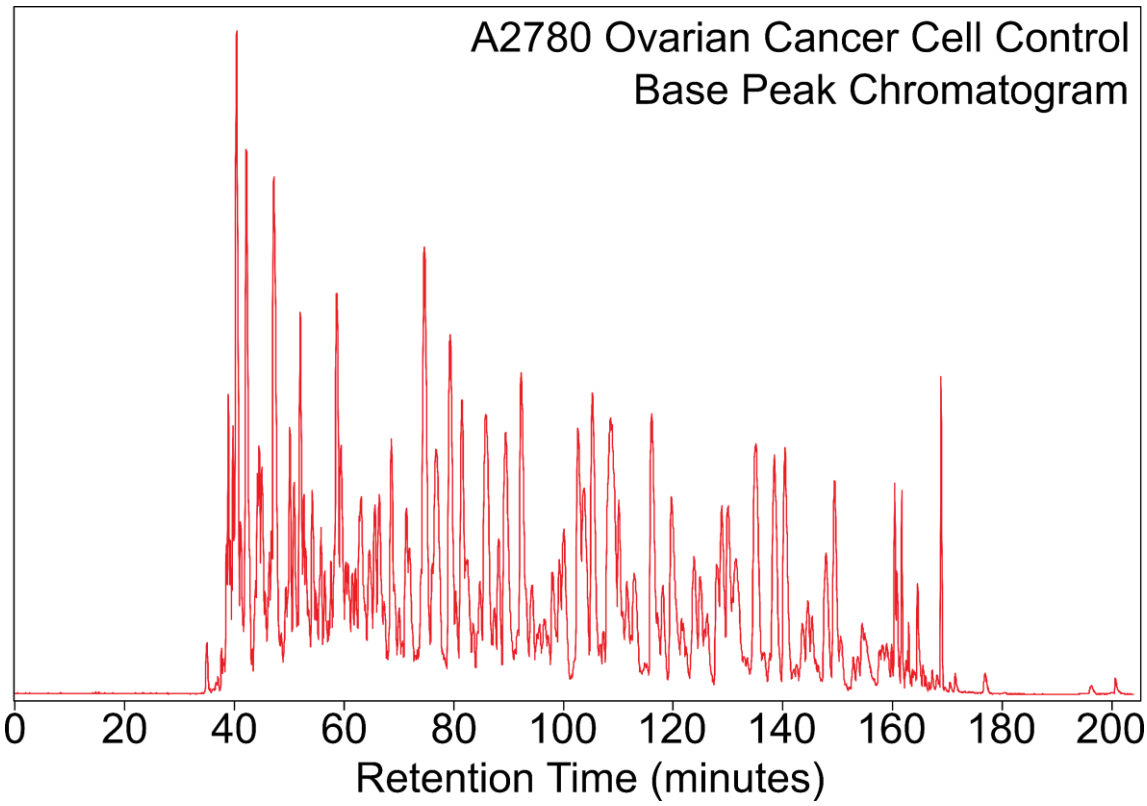


Figure S5.2 Example of base peak chromatogram of FY12-A2780 ovarian cancer cell nLC-FT-ICR CAD MS/MS.

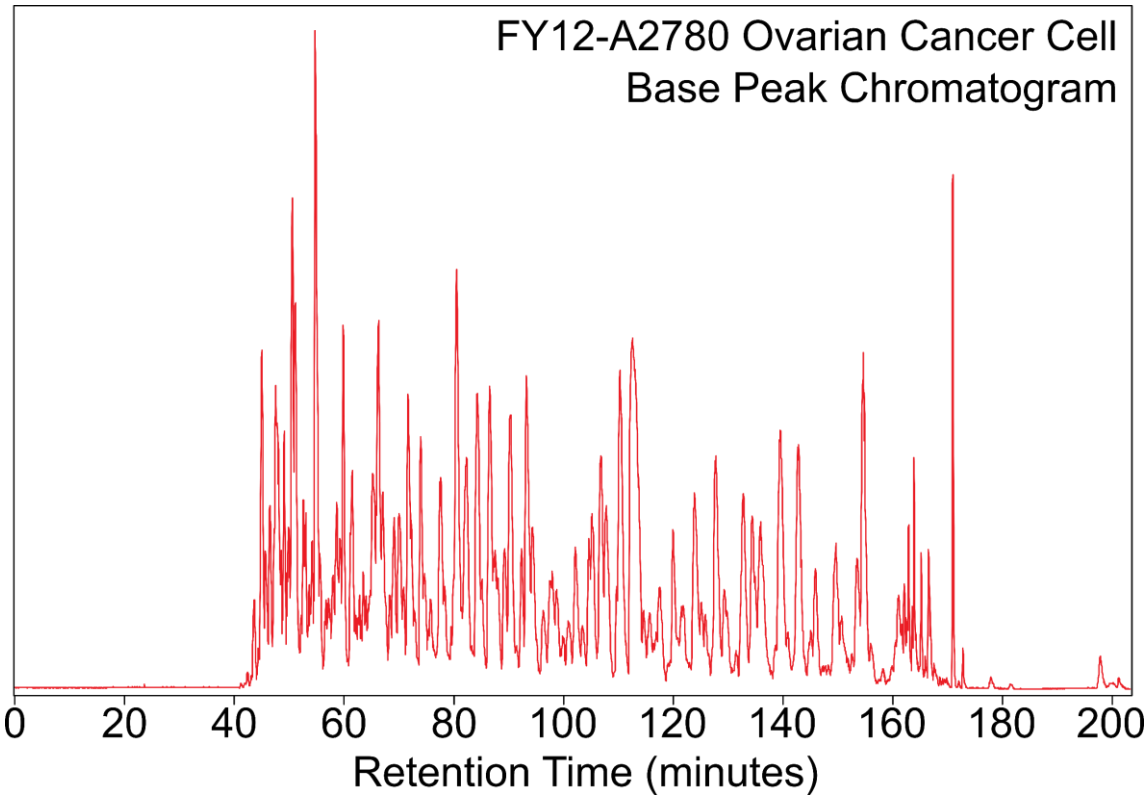


Table S5.1 T-test result for the lysozyme reference peptide (FESNFNTQATNR, 714.8365 *m/z*).

T Test between	P value
Control and Drug treated	0.34

Table S5.2 Quantification of all detected peptides in the FY12-A2780 ovarian cancer cell experiment, using the label-free approach.

Protein Description	Metabolic Pathway	Peptide Sequence	p value	Fold of Change	Average Fold of Change
Heat shock 70 kDa protein	Antigen Processing and Presentation	LKETAESVLK	2.72E-03	1.24	0.93
	Antigen Processing and Presentation	FGDPVVQSDMK	4.58E-02	1.22	
	Antigen Processing and Presentation	VEIIANDQGNR	1.97E-02	0.74	
	Antigen Processing and Presentation	DAGTIAGLNVLR	3.02E-02	1.46	
	Antigen Processing and Presentation	FEELNADLFR	7.43E-04	1.81	
Heat shock protein HSP 90-alpha	Antigen Processing and Presentation	ELISNSSDALDK	6.46E-08	2.07	2.39
	Antigen Processing and Presentation	TLTIVDTGIGMTK	1.31E-03	2.40	
Protein disulfide-isomerase	Antigen Processing and Presentation	QAGPASVPLR	7.18E-03	1.57	1.57
Alpha-enolase	Biosynthesis of Amino Acids	LMIEMDGTENK	2.76E-02	0.76	0.77
	Biosynthesis of Amino Acids	SGKYDLDFK	2.27E-02	0.80	
D-3-phosphoglycerate dehydrogenase	Biosynthesis of Amino Acids	QIPQATASMK	1.35E-03	1.76	1.11
	Biosynthesis of Amino Acids	AGTGVDNVDLEAATR	6.53E-04	0.72	
Fructose-bisphosphate aldolase	Biosynthesis of Amino Acids	ALANSLACQGK	3.97E-03	1.44	1.37
	Biosynthesis of Amino Acids	PYQYPALTPEQK	5.04E-04	1.37	
	Biosynthesis of Amino Acids	LQSIGTENTEENR	3.60E-05	0.30	
Phosphoglycerate kinase	Biosynthesis of Amino Acids	AHSSMVGVNLPQK	4.55E-03	0.84	0.83
	Biosynthesis of Amino Acids	VDFNVPMK	1.96E-02	0.75	
Phosphoglycerate mutase	Biosynthesis of Amino Acids	HYGGLTGLNK	6.42E-04	1.69	1.66
	Biosynthesis of Amino Acids	VLIAAHGNSLR	1.33E-03	1.56	
Pyruvate kinase	Biosynthesis of Amino Acids	ITLDNAYMEK	2.27E-02	1.34	1.34
S-adenosylmethionine synthase isoform	Biosynthesis of Amino Acids	FVIGGPQGDAGLTGR	1.19E-03	0.53	0.53
Triosephosphate isomerase	Biosynthesis of Amino Acids	IAVAAQNCYK	1.49E-06	7.62	2.68
	Biosynthesis of Amino Acids	SNVSDAVAQSTR	8.78E-03	0.52	

6-phosphogluconate dehydrogenase	Carbon Metabolism	SAVENCQDSWR	7.10E-04	0.76	0.76
Alpha-enolase	Carbon Metabolism	SGKYDLDFK	2.27E-02	0.80	0.77
	Carbon Metabolism	LMIEMDGTENK	2.76E-02	0.76	
D-3-phosphoglycerate dehydrogenase	Carbon Metabolism	QIPQATASMK	1.35E-03	1.76	1.11
	Carbon Metabolism	AGTGVDNVDLEAATR	3.47E-02	0.61	
Fructose-bisphosphate aldolase	Carbon Metabolism	ALANSLACQGK	3.97E-03	1.44	1.37
	Carbon Metabolism	PYQYPALTPEQK	5.04E-04	1.37	
	Carbon Metabolism	LQSIGTENTEENR	3.60E-05	0.30	
Glucose-6-phosphate isomerase	Carbon Metabolism	HFVALSTNTTK	8.13E-03	0.67	0.68
	Carbon Metabolism	TFTTQETITNAETAK	4.50E-06	0.69	
Malate dehydrogenase	Carbon Metabolism	LGVTANDVK	4.66E-02	0.71	0.80
	Carbon Metabolism	MISDAIPELK	2.66E-02	0.87	
Phosphoglycerate kinase	Carbon Metabolism	VDFNVPMK	1.96E-02	0.75	0.83
	Carbon Metabolism	AHSSMVGVNLPQK	4.55E-03	0.84	
Phosphoglycerate mutase	Carbon Metabolism	HYGGLTGLNK	6.42E-04	1.69	1.66
	Carbon Metabolism	VLIAAHGNSLR	1.33E-03	1.56	
Pyruvate kinase	Carbon Metabolism	ITLDNAYMEK	2.27E-02	1.34	1.34
Triosephosphate isomerase	Carbon Metabolism	IAVAAQNCYK	1.49E-06	7.62	2.68
	Carbon Metabolism	SNVSDAVAQSTR	8.78E-03	0.52	
14-3-3 protein epsilon	Cell Cycle	YLAEFATGNDRK	1.98E-04	0.78	0.78
14-3-3 protein gamma	Cell Cycle	YLAEVATGEK	7.84E-03	0.75	0.75
14-3-3 protein theta	Cell Cycle	SICTTVLELLDK	2.56E-03	1.76	1.04
	Cell Cycle	AVTEQGAELSNEER	4.47E-03	0.67	
14-3-3 protein zeta/delta	Cell Cycle	GIVDQSQQAYQEAFEISKK	3.04E-03	0.64	0.69
	Cell Cycle	SVTEQGAELSNEER	7.67E-04	0.70	
Proliferating cell nuclear antigen	Cell Cycle	FSASGELGNGNIK	1.57E-02	1.18	1.18
Adenosylhomocysteinase	Cys and Met Metabolism	WLNENAVEK	4.11E-02	0.60	0.65
	Cys and Met Metabolism	VAVVAGYGDVGK	4.49E-03	0.66	
L-lactate dehydrogenase	Cys and Met Metabolism	VTLTSEEEAR	1.57E-03	0.65	1.58
	Cys and Met Metabolism	LKDDEVAQLK	6.16E-04	0.76	
	Cys and Met Metabolism	VIGSGCNLDSAR	1.07E-04	1.70	
Malate dehydrogenase	Cys and Met Metabolism	LGVTANDVK	4.66E-02	0.71	0.80
	Cys and Met Metabolism	MISDAIPELK	2.66E-02	0.87	
S-adenosylmethionine synthase isoform	Cys and Met Metabolism	FVIGGPQGDAGLTGR	1.19E-03	0.53	0.53

Alpha-enolase	Glycolysis	LMIEMDGTENK	2.76E-02	0.76	0.77
	Glycolysis	SGKYDLDFK	2.27E-02	0.80	
Fructose-bisphosphate aldolase	Glycolysis	PYQYPALTPEQK	5.04E-04	1.37	1.37
	Glycolysis	LQSIGTENTEENR	3.60E-05	0.30	
	Glycolysis	ALANSLACQGK	3.97E-03	1.44	
Glucose-6-phosphate isomerase	Glycolysis	HFVALSTNTTK	8.13E-03	0.67	0.68
	Glycolysis	TFTTQETITNAETAK	4.50E-06	0.69	
L-lactate dehydrogenase	Glycolysis	VIGSGCNLDSAR	1.07E-04	1.70	1.58
	Glycolysis	VTLTSEEEAR	1.57E-03	0.65	
	Glycolysis	LKDDEVAQLK	6.16E-04	0.76	
Phosphoenolpyruvate carboxykinase	Glycolysis	TVIVTPSQR	9.85E-04	2.46	2.46
Phosphoglycerate kinase	Glycolysis	VDFNVPMK	1.96E-02	0.75	0.83
	Glycolysis	AHSSMVGVNLPQK	4.55E-03	0.84	
Phosphoglycerate mutase	Glycolysis	HYGGLTGLNK	6.42E-04	1.69	1.66
	Glycolysis	VLIAAHGNSLR	1.33E-03	1.56	
Pyruvate kinase	Glycolysis	ITLDNAYMEK	2.27E-02	1.34	1.34
Triosephosphate isomerase	Glycolysis	IAVAAQNCYK	1.49E-06	7.62	2.68
	Glycolysis	SNVSDAVAQSTR	1.35E-02	0.49	
6-phosphogluconate dehydrogenase	Pentose Phospate Metabolism	SAVENCQDSWR	7.10E-04	0.76	0.76
Fructose-bisphosphate aldolase	Pentose Phospate Metabolism	ALANSLACQGK	3.97E-03	1.44	1.37
	Pentose Phospate Metabolism	PYQYPALTPEQK	5.04E-04	1.37	
	Pentose Phospate Metabolism	LQSIGTENTEENR	3.60E-05	0.30	
Glucose-6-phosphate isomerase	Pentose Phospate Metabolism	HFVALSTNTTK	8.13E-03	0.67	0.68
	Pentose Phospate Metabolism	TFTTQETITNAETAK	4.50E-06	0.69	
Actin	Phagosome	HQGVMVGMGQK	4.11E-02	2.26	2.10
	Phagosome	GYSFTTTAER	2.51E-03	1.67	
Tubulin	Phagosome	IHFPLATYAPVISAEK	4.00E-02	1.21	3.04
	Phagosome	FPGQLNADLR	4.03E-05	8.72	
	Phagosome	TAVCDIPPR	7.37E-03	0.66	

Heat shock cognate 71 kDa protein	Protein Processing in ER	TWNDPSVQQDIK	5.12E-03	0.83	0.90
	Protein Processing in ER	DAGTIAGLNVLR	3.02E-02	1.46	
	Protein Processing in ER	FGDPVVQSDMK	4.58E-02	1.22	
	Protein Processing in ER	VEIIANDQGNR	1.97E-02	0.74	
	Protein Processing in ER	FEELNADLFR	7.43E-04	1.81	
Endoplasmin	Protein Processing in ER	ELISNASDALDK	5.91E-04	2.01	2.01
Heat shock protein HSP 90-alpha	Protein Processing in ER	ELISNSSDALDK	6.46E-08	2.07	2.39
	Protein Processing in ER	TLTIVDTGIGMTK	1.31E-03	2.40	
Protein disulfide-isomerase	Protein Processing in ER	TGEAIVDAALSALR	1.08E-02	1.67	1.65
	Protein Processing in ER	QAGPASVPLR	7.18E-03	1.57	
L-lactate dehydrogenase	Pyruvate Metabolism	VTLTSEEEAR	1.57E-03	0.65	1.58
	Pyruvate Metabolism	LKDDEVAQLK	6.16E-04	0.76	
	Pyruvate Metabolism	VIGSGCNLDSAR	1.07E-04	1.70	
Malate dehydrogenase	Pyruvate Metabolism	LGVTANDVK	4.66E-02	0.71	0.80
	Pyruvate Metabolism	MISDAIPELK	2.66E-02	0.87	
Phosphoenolpyruvate carboxykinase	Pyruvate Metabolism	TVIVTPSQR	9.85E-04	2.46	2.46
Pyruvate kinase	Pyruvate Metabolism	ITLDNAYMEK	2.27E-02	1.34	1.34
14-3-3 protein epsilon	Signalling Pathway	YLAEFATGNDRK	1.98E-04	0.78	0.78
14-3-3 protein gamma	Signalling Pathway	YLAEVATGEK	7.84E-03	0.75	0.75
14-3-3 protein theta	Signalling Pathway	SICTTVLELLDK	2.56E-03	1.76	1.04
	Signalling Pathway	AVTEQGAELSNEER	4.47E-03	0.67	
14-3-3 protein zeta/delta	Signalling Pathway	GIVDQSQQAYQEAFEISKK	3.04E-03	0.64	0.69
	Signalling Pathway	SVTEQGAELSNEER	7.67E-04	0.70	
Actin	Signalling Pathway	GYSFTTTAER	2.51E-03	1.67	2.10
	Signalling Pathway	HQGVMVGMGQK	4.11E-02	2.26	

Table S5.3 Summary of the quantified proteins in metabolic pathways of FY12-A2780 ovarian cancer cell experiment. Up-regulated proteins are shown in green and down-regulated proteins are shown in red.

Glycolysis		
Protein Description	**Average Fold of Change**	**Standard Deviation**
Alpha-enolase	0.77	0.005
Fructose-bisphosphate aldolase	1.37	0.007
Glucose-6-phosphate isomerase	0.68	0.003
L-lactate dehydrogenase	1.58	0.010
Phosphoenolpyruvate carboxykinase	2.46	0.002
Phosphoglycerate kinase	0.83	0.002
Phosphoglycerate mutase	1.66	0.004
Pyruvate kinase	1.34	0.006
Triosephosphate isomerase	2.68	0.035
Carbon Metabolism		
Protein Description	**Average Fold of Change**	**Standard Deviation**
6-phosphogluconate dehydrogenase	0.76	0.001
Alpha-enolase	0.77	0.005
D-3-phosphoglycerate dehydrogenase	1.11	0.007
Fructose-bisphosphate aldolase	1.37	0.007
Glucose-6-phosphate isomerase	0.68	0.003
Malate dehydrogenase	0.80	0.002
Phosphoglycerate kinase	0.83	0.002
Phosphoglycerate mutase	1.66	0.004
Pyruvate kinase	1.34	0.006
Triosephosphate isomerase	2.68	0.035
Signalling Pathway		
Protein Description	**Average Fold of Change**	**Standard Deviation**
14-3-3 protein epsilon	0.78	0.001
14-3-3 protein gamma	0.75	0.001
14-3-3 protein theta	1.04	0.001
14-3-3 protein zeta/delta	0.69	0.002
Actin	2.10	0.079

Cys and Met Metabolism

Protein Description	Average Fold of Change	Standard Deviation
Adenosylhomocysteinase	0.65	0.001
L-lactate dehydrogenase	1.58	0.010
Malate dehydrogenase	0.80	0.002
S-adenosylmethionine synthase isoform	0.53	0.003

Antigen Processing and Presentation

Protein Description	Average Fold of Change	Standard Deviation
Heat shock 70 kDa protein family	0.93	0.008
Heat shock protein HSP 90-alpha	2.39	0.019
Protein disulfide-isomerase	1.57	0.002

Pentose Phosphate Metabolism

Protein Description	Average Fold of Change	Standard Deviation
6-phosphogluconate dehydrogenase	0.76	0.001
Fructose-bisphosphate aldolase	1.37	0.007
Glucose-6-phosphate isomerase	0.68	0.003

Biosynthesis of Amino Acids

Protein Description	Average Fold of Change	Standard Deviation
Alpha-enolase	0.77	0.005
D-3-phosphoglycerate dehydrogenase	1.11	0.007
Fructose-bisphosphate aldolase	1.37	0.007
Phosphoglycerate kinase	0.83	0.002
Phosphoglycerate mutase	1.66	0.004
Pyruvate kinase	1.34	0.006
S-adenosylmethionine synthase isoform	0.53	0.003
Triosephosphate isomerase	2.68	0.035

Protein Processing in Endoplasmic Reticulum

Protein Description	Average Fold of Change	Standard Deviation
Endoplasmin	2.01	0.009
Heat shock 70 kDa protein family	0.93	0.008
Heat shock protein HSP 90-alpha	2.39	0.019
Protein disulfide-isomerase	1.57	0.002

Cell Cycle

Protein Description	Average Fold of Change	Standard Deviation
14-3-3 protein epsilon	0.78	0.001
14-3-3 protein gamma	0.75	0.001
14-3-3 protein theta	1.04	0.001
14-3-3 protein zeta/delta	0.69	0.002
Proliferating cell nuclear antigen	1.18	0.001

Pyruvate Metabolism

Protein Description	Average Fold of Change	Standard Deviation
L-lactate dehydrogenase	1.58	0.010
Malate dehydrogenase	0.80	0.002
Phosphoenolpyruvate carboxykinase	2.46	0.002
Pyruvate kinase	1.34	0.006

Phagosome

Protein Description	Average Fold of Change	Standard Deviation
Actin	2.10	0.079
Tubulin	3.04	0.010

Figure S5.3 Example of base peak chromatogram of *S. aureus* bacterial cell control nLC-FT-ICR CAD MS/MS.

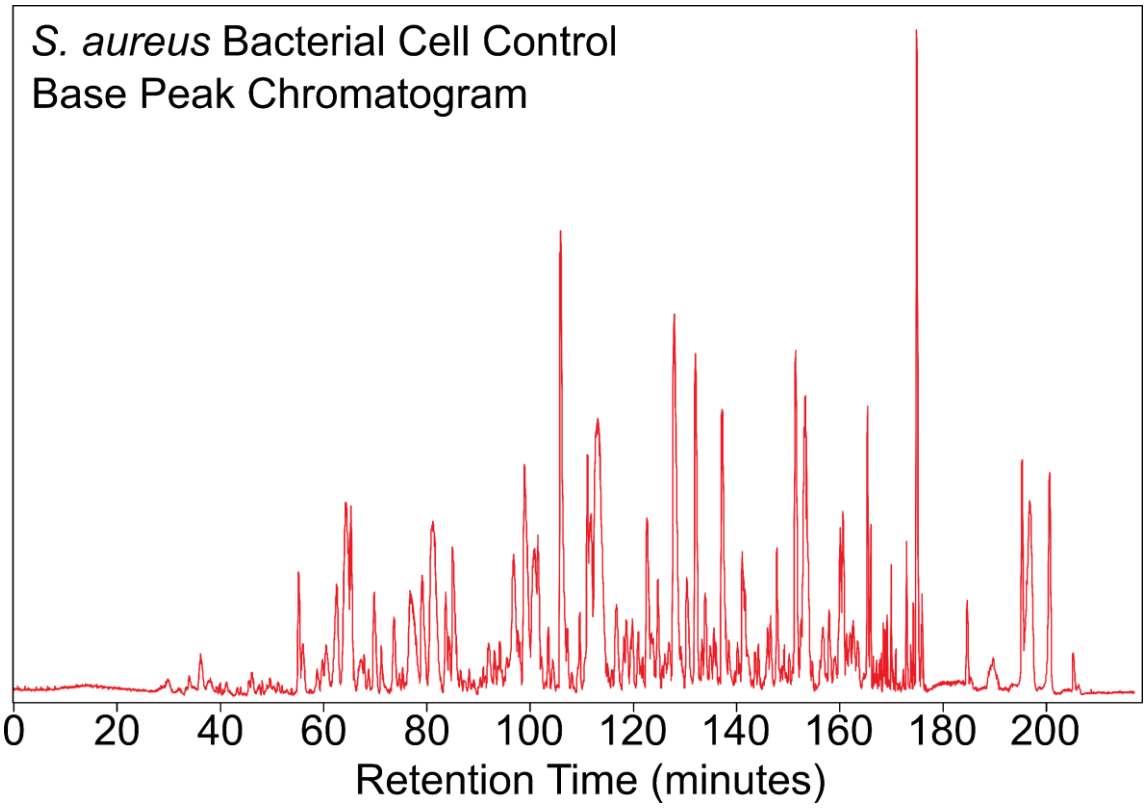


Figure S5.4 Example of base peak chromatogram of CF281-*S. aureus* bacterial cell nLC-FT-ICR CAD MS/MS.

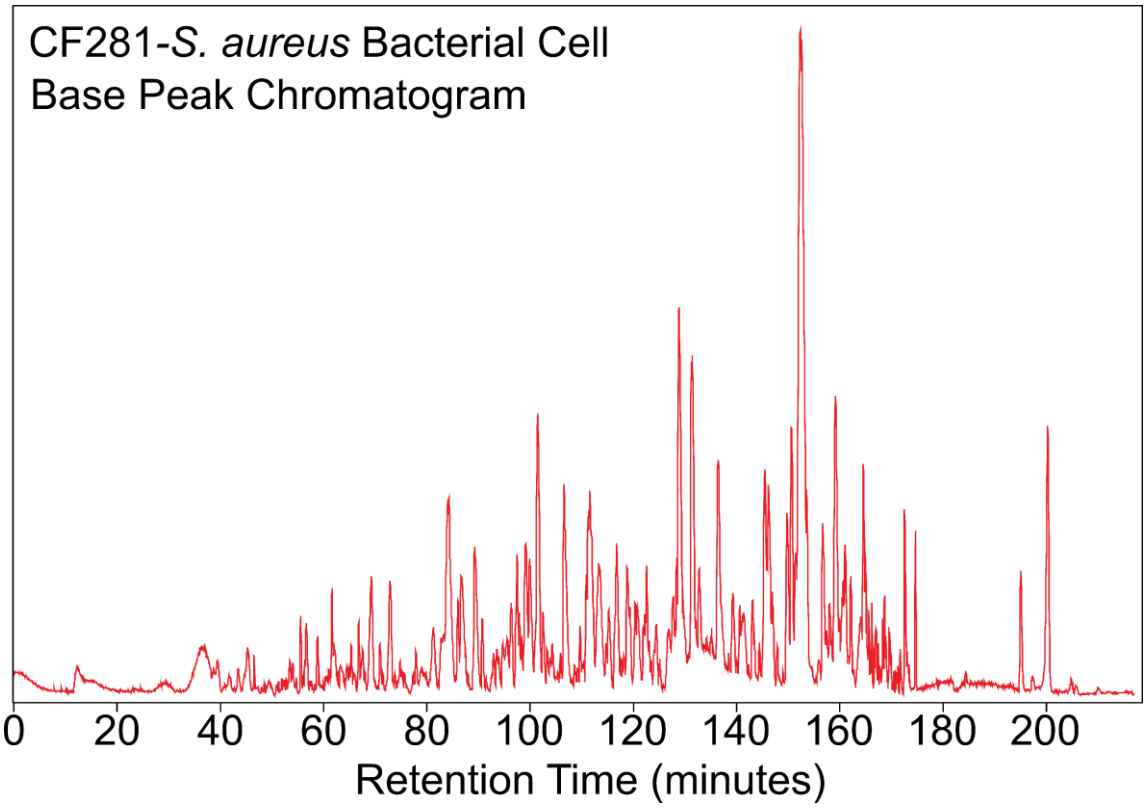


Table S5.4 T-test result for the lysozyme reference peptide (NTDGSTDYGILQINSR, 877.4212 *m/z*).

T Test between	P value
Control and Drug Treated (1 MIC)	0.12

Table S5.5 Quantification of all detected peptides for the CF281-*S. aureus* bacterial cell experiment using the label-free approach.

Protein Description	Peptide Sequence	p value	Fold of Change	Average Fold of Change
30S ribosomal protein S2	WLGGILTNYK	1.43E-02	2.10	2.10
30S ribosomal protein S7	TMEDRLANEILDAANNTGGAVK	9.95E-04	3.57	3.57
50S ribosomal protein L13	LFVYGGAEHPHAAQQPENYELRG	2.44E-05	2.33	2.33
50S ribosomal protein L14	KTANIGDVIVCTVK	7.59E-04	2.73	2.73
50S ribosomal protein L2	NMTSLDFAEITK	1.86E-03	3.96	3.96
50S ribosomal protein L25	TVEVPVQLVGEAVGAK	7.74E-05	12.42	12.42
50S ribosomal protein L5	VLDNAVEELELITGQKPLVTK	9.47E-04	4.91	4.91
Acyl carrier protein	MENFDKVKDIIVDR	7.56E-03	6.18	6.18
Alcohol dehydrogenase	IVAFDINDDKLAFAK	3.93E-04	10.02	6.52
	VVAVGLPVDKMNLDIPR	1.05E-03	4.18	
	LVLDGIEVVGSLVGTR	2.91E-02	7.11	
Alkyl hydroperoxide reductase protein C	NFDVLDEATGLAQR	1.56E-04	8.44	8.44
ATP synthase subunit b	QQQEQIIHEANVR	5.52E-03	2.40	2.40
Elongation factor G	VMTDPYVGKLTFFR	1.42E-04	7.89	7.89
Elongation factor Ts	ALTETDGDIDKAIDYLREK	3.16E-02	2.57	2.57
Enolase	AAADLLGQPLYK	1.10E-05	8.55	8.55
Formate acetyltransferase	ETLIDAMEHPEEYPQLTIR	2.42E-04	9.52	7.63
	AACEAYGYELDEETEKIFTDYR	4.94E-03	3.55	
	ETLLDAMEHPEEYPQLTIR	2.42E-04	9.52	
Immunodominant staphylococcal antigen A	LSNGNTAGATGSSAAQIMAQR	6.92E-04	11.85	11.85
Immunoglobulin-binding protein sbi	GAIDQTVLTVLGSGSK	7.64E-04	3.78	3.78
L-lactate dehydrogenase	TLKDIMAEAEELK	3.69E-03	2.83	2.83
L-threonine dehydratase catabolic TdcB	LLGIDATIVMPETAPIAK	3.17E-04	7.09	7.09
Probable DNA-directed RNA polymerase subunit delta	FLNVGENLWGLRDWYSVDDIEEK	4.76E-04	10.24	10.24
Pyrimidine-nucleoside phosphorylase	TLEDAEALAHAMVR	1.51E-02	2.24	2.50
	NTMAIISDMNQPLGR	3.48E-03	2.71	
Pyruvate dehydrogenase E1 component beta subunit	TVQPLDVDTIVASVEK	5.38E-04	6.53	6.53
	TVQPIDVDTIVASVEK	5.03E-04	7.09	
Pyruvate kinase	STDALLNNAVATAVETGR	6.71E-05	7.92	7.55
	AEASDVANAIYDGTDAVMLSGETAAGLYPEEAVK	1.02E-02	5.07	
Succinyl-diaminopimelate desuccinylase	IATVENPELSFASK	9.97E-06	8.56	8.56
Thiol peroxidase	LISVVPSIDTGVCDQQTR	2.44E-02	2.43	2.43
Triosephosphate isomerase	SSTSEDANEMCAFVR	5.96E-06	8.32	7.16
	SVVIAYEPIWAIGTGK	6.71E-04	7.18	
	QTIADLSSKEVSEATR	1.40E-04	3.21	

Chapter 6

Conclusions and Future Work

The work presented in this thesis has illustrated new progress in studies of metallodrugs and their interaction with biomolecules using Fourier Transform-Ion Cyclotron Resonance Mass Spectrometry. This chapter aims to introduce, summarise, and outline the achievements and conclusions of each of the previous chapters.

Chapter 2 Metallodrug-Peptide Interactions

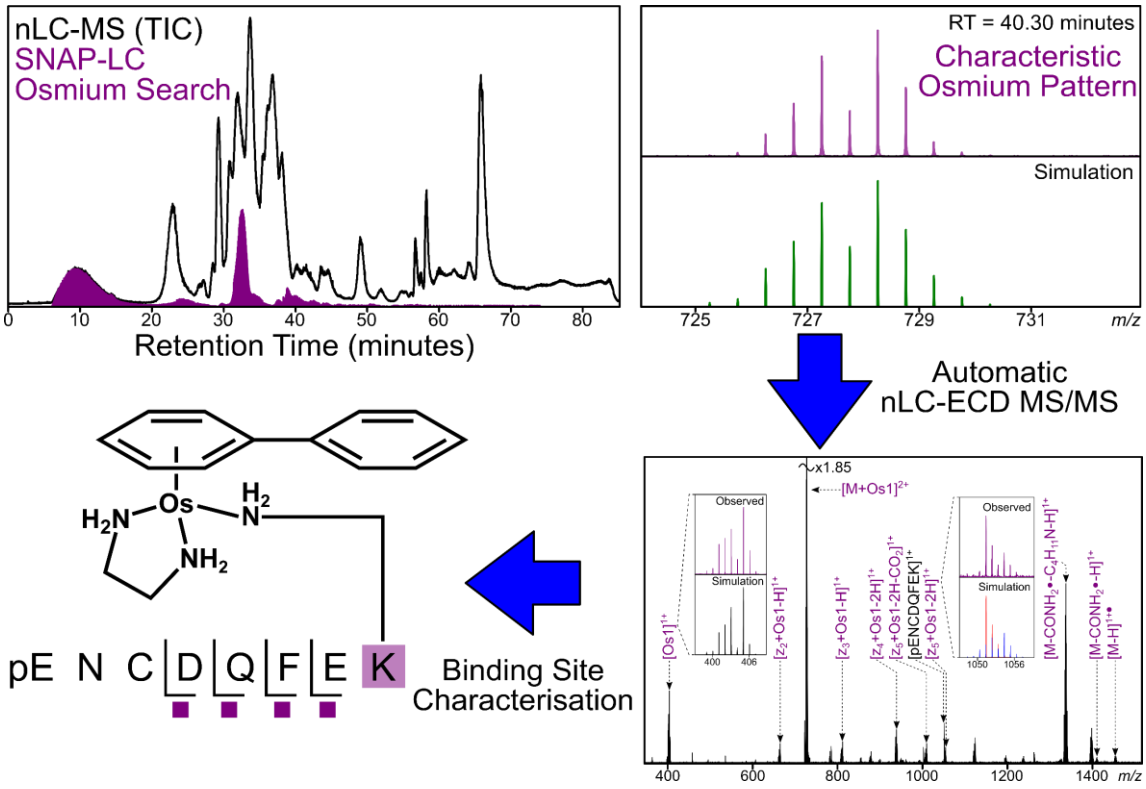


Figure 6.1 Summarising graphic for Chapter 2

The novel osmium-(II)-based anticancer complex, $[(\eta^6\text{-bip})Os(en)Cl]PF_6$ (Os1-Cl), was shown to interact with DNA with a different mechanism of action to cisplatin, which could lead to new ways of potentially attacking/killing cancerous cells more effectively with less undesired side effects. Interactions between Os1-Cl and proteins/peptides were also studied to discover the possibility of possessing bimodal mechanisms of action.

Tryptic digested bovine serum albumin (BSA) was used as a model protein and a series of osmiated BSA peptides were observed. The osmium complex was found to dissociate readily with low collision energy of 5 V, producing the unmodified peptide, suggesting the binding interaction between the osmium complex and the target peptide is weaker than the protein backbone. Electron capture dissociation tandem mass spectrometry (ECD MS/MS) successfully preserved the osmium modification upon fragmentation and more binding information was revealed. Os1-Cl was found to interact with the side chain of a lysine residue on the peptide KFWG<u>K</u> as well as a weak π-π interaction between the biphenyl ligand and the aromatic side chain of tryptophan.

nLC-MS and nLC-MS/MS using a reversed-phase (RP) column were also attempted to provide more separation between the unmodified and osmiated peptides. An isotopic prediction algorithm (SNAP-LC algorithm) was used to pick out species containing a characteristic osmium isotope pattern and generate a prefer/inclusion list for targeted nLC-ECD MS/MS analysis. Side chains of glutamic acid and lysine, as well as the N- and C- termini were revealed to be possible osmium binding sites.

This work shows FT-ICR MS is particularly suitable for studying metallodrug-biomolecule interactions with the wide array of possible fragmentation techniques and capable of coupling to different hyphenated techniques. The ultra-high resolution and mass accuracy provided by FT-ICR MS not only allow confident characterisations but also provide important knowledge for new anticancer metallodrug designs. The methodology developed can effectively identify metal-bound species by SNAP-LC peak picking algorithm, accurately identify the characteristic isotope pattern of the transition metals. Together with a wide variety of fragmentation techniques, this workflow can be applied to more complex systems, *e.g.* metallodrug-intact protein interaction analysis, or even metallodrug-whole cell interaction analysis, in order to fully characterise the properties of these novel metallodrug complexes.

Chapter 3 Reduction of Disulfide Bonds in Proteins for Enhanced MS and MS/MS Capabilities

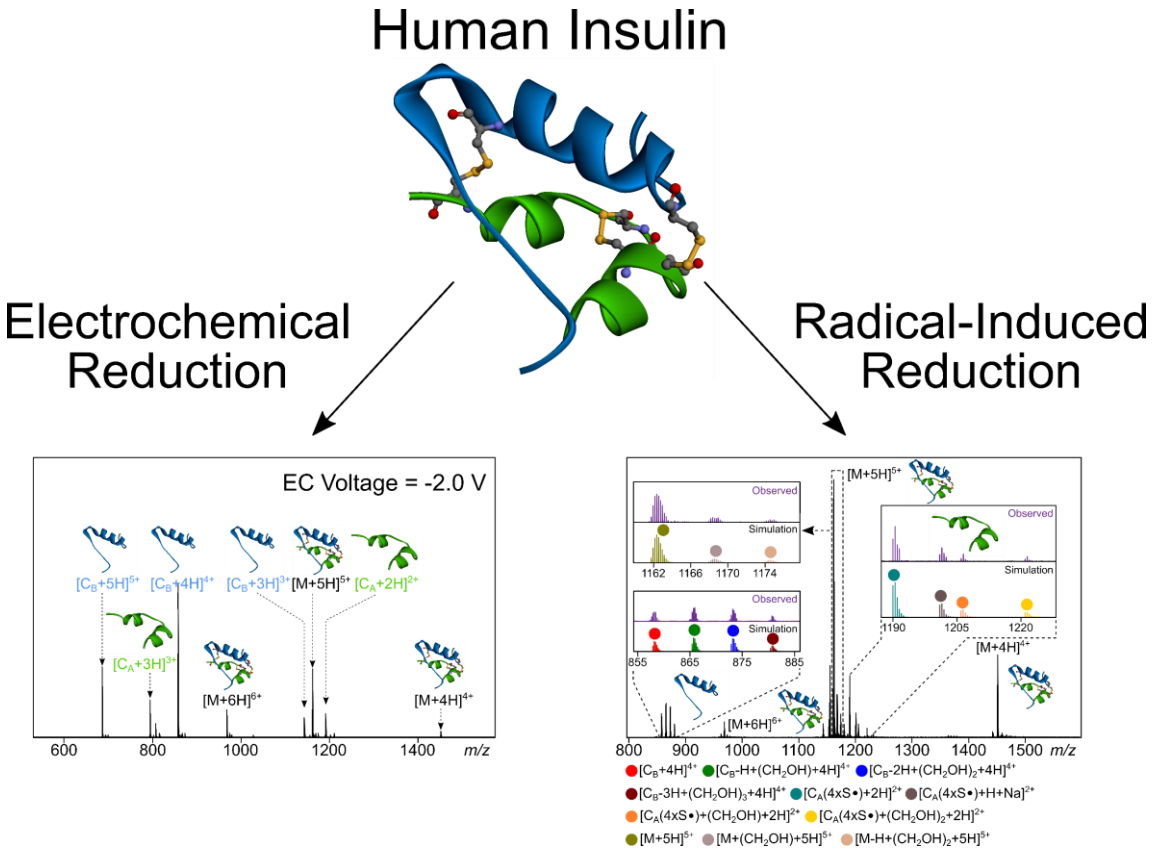


Figure 6.2 Summarising graphic for Chapter 3

Chapter 3 focuses on alternative methods for disulfide bond reduction. Traditional reduction methods use chemical reducing reagents, such as dithiothreitol (DTT) to reduce disulfide bonds. However, use of chemical-based reducing reagents often requires harsh reaction conditions, such as an excessive amount of concentrated reagents and specific pH adjustments. Labile modifications, such as metallodrug binding can easily be lost/altered during the chemical reduction, results in misleading conclusions. Therefore, the alternative reduction methods; electrochemical reduction and radical-induced reduction were investigated in this chapter.

The electrochemical reduction process was monitored using human lysozyme, bovine ribonuclease A and human insulin. The reduction process was shown to be a sequential process and could be used to selectively reduced certain disulfide bond in

stages. The reduced protein species were shown to be more effectively ionised during the electrospray ionisation (ESI) process and attained higher charge states than non-reduced species. It is believed that reduced species have more open protein structure with more sites exposed to charge carriers, such as protons. The electrochemical voltage to reduce fully a disulfide-bonded protein was also found to be protein-dependent, which can be measured using a statistical approach, especially with large intra-disulfide bonded proteins, where the change in mass is often minimal compared to the mass of the protein. Insulin was used as a model inter-disulfide bonded protein, two separated polypeptide chains were observed upon reduction. The ion abundance of the polypeptide chains was found to be the dominant peaks in the ESI-MS, overtaking the intact insulin species. The reduced species were shown to undergo fragmentation more readily in subsequent MS/MS experiments. An increase in cleavage coverage of over 7-fold was observed with human lysozyme compared to the non-reduced species upon ECD MS/MS.

The performance of radical-induced reduction was monitored using human insulin as a model protein. The two separated polypeptide chains were observed to increase in intensity with the irradiation time, suggesting the reduction process was occurring in the solution phase. However, different to electrochemical reduction, the intact protein species remained the dominant species in the MS spectrum. Limited fragmentation of intact insulin was achieved using collisionally activated dissociation tandem mass spectrometry (CAD MS/MS), cleavage coverage of only 39% was observed, while ECD MS/MS provided a coverage of 51%, an increase believed to be contributed by the disulfide bond preferentiality of ECD MS/MS. Upon radical-induced reduction, a promising 79% cleavage coverage with ECD MS/MS of the chain B of insulin was observed.

Both alternative reduction methods have proved to be effective tools for top-down analysis and can greatly assist protein identification and characterisation, which is a key research area in proteomics studies, especially those containing chemically sensitive modifications, such as metallodrugs. These alternative reduction methods can be applied to the metallodrug-biomolecules system, to avoid the use of chemical

reducing reagents like DTT that could disrupt the labile metallodrug-modifications in the future to achieve a more comprehensive characterisation of these novel metal complexes.

Chapter 4 Organoiridium Photosensitisers Induce Specific Oxidative Attack on Proteins within Cancer Cells

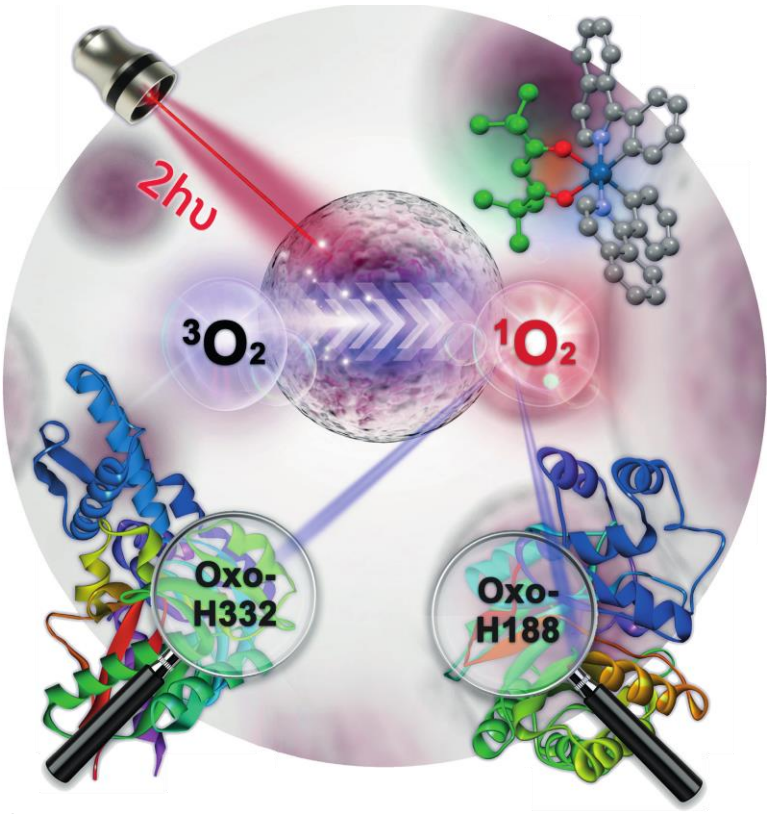


Figure 6.3 Summarising graphic for Chapter 4 (Reproduced from Zhang *et. al.* 2017 with permission).[1,2]

Six iridium (III)-based photosensitisers were designed to release cytotoxic singlet oxygen when irradiated with visible blue light (463 nm). The mechanisms of action of these novel photosensitisers on a model peptide, [Lys^3]-Bombesin were studied using FT-ICR MS and ECD MS/MS. Four of the six photosensitisers showed the ability to not only oxidise, but also chemically bind to the target peptide, which displayed the possibility of a dual mechanisms of action against target biomolecules. A series of oxidation products were observed, and by employing ECD MS/MS, oxidation sites were found to be methionine, histidine, and tryptophan. A specific form of tryptophan, 3-hydroxykynurenine, was identified and provided solid evidence for the production of singlet oxygen. Histidine was found to oxidise to 2-oxo-histidine, which is also a common oxidation product of histidine upon singlet oxygen oxidation.

Iridium complex PZ6 showed no toxicity towards A549 3D multicellular tumour spheroids (MCTSs) in the dark, and no toxicity towards MRC-5 MCTSs in either the dark or upon irradiation. However, the IC_{50} values for PZ6 towards A549 spheroids upon 1P and 2P irradiation were extremely low, 1.0 μM and 0.23 μM respectively, showing that PZ6 is a potent anticancer metallodrug with high selectivity to cancerous cells.

The effect of PZ6 on A549 lung cancer cells was studied by nLC-FT-ICR CAD MS/MS and the proteins were quantified using a label-free approach. Specific histidine residues in the proteins heat shock protein 70 kDa and aldose reductase were found to be oxidised and potentially disrupt their cellular functions in cancer cells. The combination of oxidative stress induced by photoactivation of PZ6 together with the malfunction of mitochondria in cancer cells leads to the increased use of glucose to generate energy and is consistent with the observation of the increase in quantity of the nine unique enzymes involved in the glycolytic pathway.

Most of the protein identified were the abundant proteins in cells, that could be due to the single-dimensional liquid chromatography experimental setup, causing co-elution of species. Ions with lower intensities were supressed by the ions with higher intensities. The low intensity species often could not be identified because of the lack of fragmentation/effective fragmentation of these species. To improve with the experimental setup, a multi-dimensional liquid chromatography system could be used, with the extra dimensional of separation (orthogonal to the first dimension, *e.g.* reverse phase column followed by strong cation exchange column), the co-eluting species have a higher probability to be separated and effectively fragmented.

Chapter 5 Metallodrug-Proteomics Pathway Analysis using nLC-FT-ICR MS

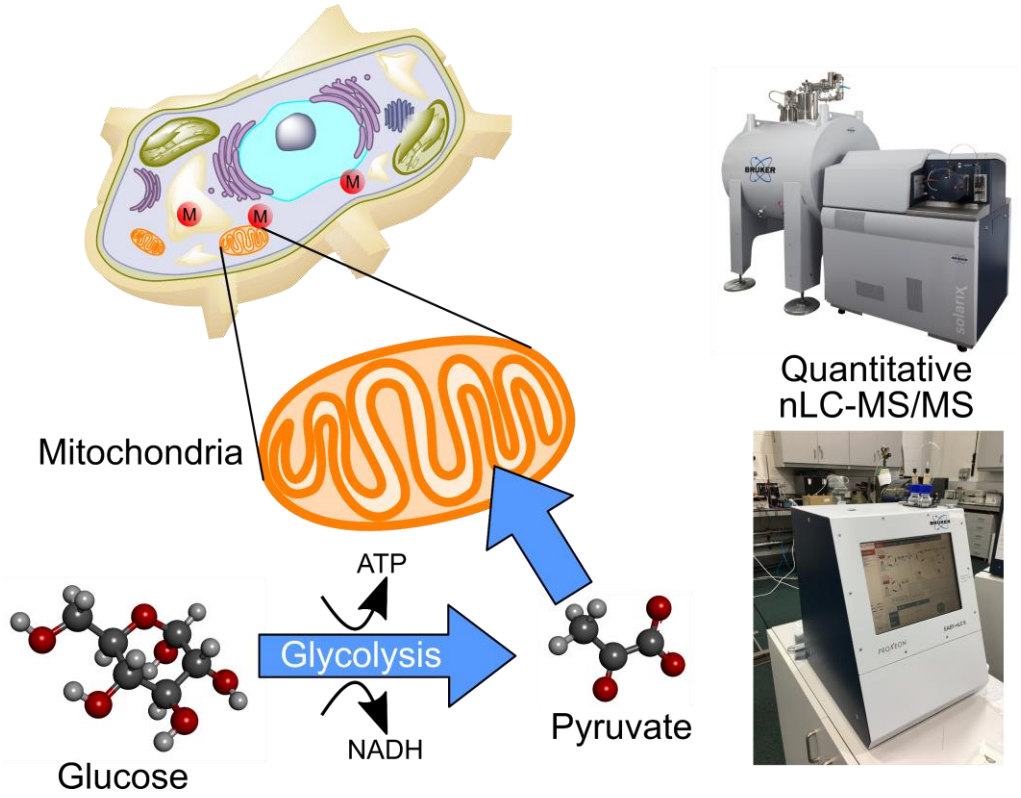


Figure 6.4 Summarising graphic for Chapter 5

Quantitative nLC-FT-ICR MS/MS was employed to study the proteomic effects of an osmium-based anticancer drug (FY12) against A2780 ovarian cancer cells and an iridium-based antimicrobial complex (CF281) against *Staphylococcus aureus (S. aureus)*.

nLC-FT-ICR CAD MS/MS was used to conduct pathway analysis of FY12 on A2780 ovarian cancer cells. The proteins were identified by matching the fragmentation pattern against the Mascot database and quantification of the identified proteins was measured using a label-free approach with a lysozyme tryptic digested peptide (FESNFNTQATNR, 714.8365 *m/z*). FY12 has shown significant effects on multiple metabolic pathways and different cellular functions. Nine unique proteins along the glycolysis pathway showed a significant change in quantity and pyruvate was selectively accumulated for extra energy production. In addition, quantities of 14-3-3 isoforms were also perturbed and believed to be the primary MoA against the ovarian

cancer cells. Different stress responses, including the up-regulation of heat shock proteins and the down-regulation of amino acid synthesis which are believed to be a secondary response from the cancer cells, were observed.

The MoA of the iridium metalloantibiotic CF281 in *S. aureus* was investigated using a similar methodology and quantified with another lysozyme tryptic digested peptide (NTDGSTDYGILQINSR, 877.4212 *m/z*). 81 unique proteins were detected three times or more out of the replicates, and 26 of those were shown to have significant up-regulated quantity change with p value less than 0.05 in a two-tailed t-test. The bacterial cells were shown to be in a very stressful chemical environment when they were treated with CF281 and triggered a series of secondary responses, including extra protein/DNA biosynthesis and energy production. CF281 also found to be capable of inducing oxidative stress inside the bacteria and trigger cell death, different to the traditional treatments for *S. aureus*, which focus on the disruption of cell wall formation by inhibiting the formation of peptidoglycan cross linkages that provide rigidity and strength for the cell wall.

FT-ICR MS has proved to be a very versatile analytical technique to study the interactions between biomolecules and metallodrugs in this thesis. The ultra-high resolution and mass accuracy of FT-ICR MS allow unambiguous assignments in mass spectra and provide important information about the mechanisms of action of the metallodrugs. FT-ICR MS together with hyphenated techniques allow a more detailed analysis in complex systems, such as cell lysate. SNAP-LC has also been shown to pick out metal-modified species effectively and precisely in mass spectra, and localisation of the metallodrug can be achieved *via* various fragmentation techniques. Constant improvement in MS techniques and instrumentations will help with the speed of data acquisition and accuracy of acquired data. A truly data independent acquisition method, two-dimensional mass spectrometry (2DMS) is currently under rapid development and allow full characterisation of metallodrug-biomolecules interaction. Analytical procedures mentioned in the thesis will hopefully become more automatic and high throughput in the future, which can help inform better drug designs, elucidate their mechanisms of actions and application to different types of life-threatening diseases.

References

(1) Zhang, P.; Chiu, C. K. C.; Huang, H.; Lam, Y. P. Y.; Habtemariam, A.; Malcomson, T.; Paterson, M. J.; Clarkson, G. J.; O'Connor, P. B.; Chao, H.; et al. Organoiridium Photosensitizers Induce Specific Oxidative Attack on Proteins within Cancer Cells. *Angew. Chemie Int. Ed.* **2017**, 14898–14902.

(2) Zhang, P.; Chiu, C. K. C.; Huang, H.; Lam, Y. P. Y.; Habtemariam, A.; Malcomson, T.; Paterson, M. J.; Clarkson, G. J.; O'Connor, P. B.; Chao, H.; et al. Inside Cover: Organoiridium Photosensitizers Induce Specific Oxidative Attack on Proteins within Cancer Cells (Angew. Chem. Int. Ed. 47/2017). *Angew. Chemie Int. Ed.* **2017**, *56* (47), 14774–14774.